



**UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**



VICTORIA GABRIELLA REBOUÇAS PEREIRA

**UMA META-ANÁLISE DA EFETIVIDADE DE MODELOS
DE APRENDIZADO POR TRANSFERÊNCIA EM
DIFERENTES CONTEXTOS EDUCACIONAIS**

MOSSORÓ

2024

VICTORIA GABRIELLA REBOUÇAS PEREIRA

**UMA META-ANÁLISE DA EFETIVIDADE DE MODELOS
DE APRENDIZADO POR TRANSFERÊNCIA EM
DIFERENTES CONTEXTOS EDUCACIONAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Lenardo Chaves e Silva - UFRSA

Coorientador: Prof. Dr. Álvaro Álvares de Carvalho César Sobrinho - UFRPE

MOSSORÓ

2024

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

PPere Pereira, Victoria.
iram UMA META-ANÁLISE DA EFETIVIDADE DE MODELOS DE
APRENDIZADO POR TRANSFERÊNCIA EM DIFERENTES
CONTEXTOS EDUCACIONAIS / Victoria Pereira. - 2024.
205 f. : il.

Orientador: Lenardo Silva.
Coorientador: Álvaro Sobrinho.
Dissertação (Mestrado) - Universidade Federal
Rural do Semi-árido, Programa de Pós-graduação em
--Selecione um Curso ou Programa--, 2024.

1. Aprendizagem por Transferência. 2. Educação.
3. Revisão Sistemática da Literatura. 4. Meta-
análise. I. Silva, Lenardo , orient. II.
Sobrinho, Álvaro, co-orient. III. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo) autor(a).
Biblioteca Campus Mossoró / Setor de Informação e Referência
Bibliotecária: Keina Cristina Santos Sousa e Silva
CRB: 15/120

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

VICTORIA GABRIELLA REBOUÇAS PEREIRA

**UMA META-ANÁLISE DA EFETIVIDADE DE MODELOS
DE APRENDIZADO POR TRANSFERÊNCIA EM
DIFERENTES CONTEXTOS EDUCACIONAIS**

Dissertação apresentada ao Programa de
Pós-Graduação em Ciência da Computação para
a obtenção do título de Mestre em Ciência da
Computação.

Aprovada em: 30/08/2024

BANCA EXAMINADORA

Prof. Dr. Lenardo Chaves e Silva - UFERSA
Presidente

Prof. Dr. Álvaro Álvares de Carvalho César Sobrinho
- UFAPE
Primeiro Membro

Prof. Dr. Patrícia Freitas Campos de Vasconcelos-
UFC
Segundo Membro

Prof. Dr. Sebastião Emidio Alves Filho- UERN
Terceiro Membro

À minha amada família, minha profunda gratidão pelo apoio incansável ao longo desta jornada.

AGRADECIMENTOS

Agradeço a Deus por todas as conquistas que alcancei até agora e por toda a força nos dias difíceis durante esta jornada. Sou grata por sua inspiração nas decisões que me trouxeram até este momento, superando todos os desafios. Meu Deus, agradeço pelos teus planos para a minha vida, pois eles sempre se mostram maiores do que meus próprios sonhos!

Aos meus amados pais, Vicente Pereira e Maria das Graças, expresso minha mais profunda gratidão por todo o amor, carinho, compreensão, apoio, paciência e os valiosos ensinamentos que vocês generosamente dedicaram a mim ao longo de toda a minha jornada. Desde os primeiros passos até os desafios mais recentes, vocês estiveram ao meu lado, guiando-me com sabedoria e nutrindo meu crescimento em todos os aspectos da vida. Agradeço também por todos os esforços incansáveis que vocês realizaram para que eu pudesse alcançar este momento. Cada sacrifício feito em prol do meu bem-estar e sucesso não passou despercebido, e sou profundamente grata por todo o apoio inabalável que sempre me proporcionaram. Vocês são os pilares da minha vida, a fonte de inspiração, e a razão pela qual estou onde estou hoje.

Queridos irmãos, Vilma Grazielle, Virginia Graziella, Victor Gabriel e Nara Livianne, neste momento, quero expressar o quanto cada um de vocês é importante para mim. Que possamos continuar caminhando lado a lado, fortalecendo nossa relação e celebrando cada conquista juntos. Saibam que estarei sempre aqui para apoiá-los e compartilhar cada momento importante de suas vidas. Nossa união é um tesouro que valorizo profundamente.

Ao meu noivo, Henrique Augusto, sua presença ao meu lado em cada momento da minha vida, especialmente durante esta jornada acadêmica, é um presente verdadeiramente inestimável. Agradeço por ser meu apoio constante, por me incentivar e por compartilhar tantos momentos preciosos enquanto eu me dedicava a esta pesquisa. Sua presença tornou os desafios mais leves e os momentos de celebração ainda mais especiais. Sou profundamente grata por ter você ao meu lado, não apenas como meu noivo, mas como meu companheiro de vida, pronto para enfrentar todas as aventuras que o futuro nos reserva juntos.

A todas as minhas queridas sobrinhas, Maria Ysabell, Ana Valentina, Ana Liz, Isabela Pereira, e ao meu sobrinho Luiz Gabriel, gostaria de expressar minha profunda

gratidão pelo carinho, risadas e motivação diária que vocês me proporcionaram. Suas presenças foram como luzes em meu caminho, lembrando-me constantemente de nunca desistir e seguir em frente com determinação. Cada momento compartilhado ao lado de vocês foi uma parte especial e significativa desta jornada, e sou imensamente grata por tê-los ao meu lado.

Aos meus queridos cunhados Mario Jeffson, Gustavo Nepomuceno, Juberto Moreira e minha cunhada Kaliane Fernandes, gostaria de expressar minha sincera gratidão por todo o apoio, compreensão e incentivo que generosamente me proporcionaram ao longo de todas as fases desta jornada. Seu apoio inabalável e palavras de encorajamento foram fundamentais para me manter firme e determinado, e por isso, deixo minha mais profunda gratidão. Obrigado por sempre acreditarem em mim e por estarem ao meu lado, tanto nesta como em outras ocasiões da minha vida. Vocês são uma parte preciosa do meu apoio e da minha jornada, e sou imensamente grato por tê-los como parte da minha família.

Aos meus amados avós maternos, Maria Elsa Rebouças Maia (IN MEMORIAN) e Hélio Rebouças Maia, que estiveram ao meu lado desde o início e sempre celebraram minhas conquistas, dedico este momento especial. Aos meus queridos avós paternos, Maria Neném de Oliveira e Carlos Pereira de Oliveira (IN MEMORIAN), embora agora estejam ao lado de Deus, sinto que compartilham comigo esta vitória de alguma forma. À minha bisavó Daura Rebouças Maia (IN MEMORIAN), que também partiu para junto de Deus, mas cuja influência foi fundamental para a realização do meu sonho, expresso minha mais profunda gratidão. Agradeço por todos os sábios conselhos, apoio incondicional, carinho afetuoso e compreensão que sempre me proporcionaram. Seus ensinamentos e amor deixaram uma marca indelével em minha vida, e sou eternamente grato por tê-los como exemplos.

Aos meus queridos padrinhos, Maria Gorete e Eleilton Juvenal, dedico minha mais profunda gratidão por todo o apoio, incentivo e orientação que generosamente me ofereceram ao longo desta jornada e em todas as fases da minha vida. Suas palavras de encorajamento e sua presença constante foram verdadeiros pilares para mim, inspirando-me a nunca desistir e a perseverar diante dos desafios. Agradeço por estarem sempre ao meu lado, prontos para oferecer seu apoio incondicional e sua sabedoria. Seu amor e amizade são inestimáveis para mim, e sou imensamente grata por tê-los como padrinhos, uma bênção que valorizo profundamente.

Ao estimado professor Dr. Leonardo Chaves e Silva, gostaria de expressar minha sincera gratidão por aceitar conduzir e orientar meu projeto acadêmico. Sua orientação, apoio e confiança foram fundamentais em cada etapa dessa jornada. Sou profundamente grata pelo incentivo constante e pela dedicação incansável que você dedicou ao meu projeto. Você investiu inúmeras horas para esclarecer minhas dúvidas, apontar o caminho certo e oferecer suporte quando mais precisei. Sua atenção e encorajamento foram essenciais para manter minha motivação durante todo o processo de pesquisa. Além disso, seus vastos conhecimentos contribuíram significativamente para o resultado final deste trabalho. Mais uma vez, obrigada por sua orientação exemplar e por ser uma fonte inestimável de inspiração e aprendizado. Sua influência deixou uma marca indelével em minha jornada acadêmica e sou imensamente grata por ter tido a oportunidade de aprender com você.

Ao estimado Coorientador, Prof. Dr. Álvaro Álvares de Carvalho César Sobrinho, gostaria de expressar minha profunda gratidão por aceitar coorientar minha pesquisa. Suas valiosas orientações e insights foram de fundamental importância ao longo de todo o processo. Agradeço especialmente pela paciência, sabedoria, dedicação e atenção que você dedicou a este projeto. Sua orientação foi essencial para superar desafios e alcançar resultados satisfatórios. Saiba que sua contribuição foi crucial para o sucesso desta jornada acadêmica. Sua presença e apoio foram verdadeiros pilares neste caminho. Sou imensamente grata por ter tido a oportunidade de aprender com você e por sua dedicação em auxiliar no meu desenvolvimento acadêmico e profissional.

Aos professores da banca examinadora da minha pesquisa, expresso minha sincera gratidão. Agradeço por todas as valiosas contribuições fornecidas durante a avaliação do meu trabalho. Seus insights e análises externas não apenas enriqueceram minha pesquisa, mas também abriram novas perspectivas de investigação, destacando a importância do diálogo e da colaboração na construção do conhecimento acadêmico.

Ao Programa de Pós-Graduação em Ciência da Computação (PPgCC) da UERN/UFERSA, gostaria de expressar minha profunda gratidão. Agradeço ao corpo docente, à direção e à administração por proporcionarem um ambiente tão propício ao aprendizado e ao compartilhamento de conhecimento. Sem o apoio e a estrutura fornecidos por esta instituição, minha jornada acadêmica não teria sido possível. Sou imensamente grata por todo o suporte e pelas oportunidades oferecidas.

Por fim, gostaria de expressar minha sincera gratidão a todos que, de alguma forma, contribuíram para a conclusão desta importante etapa. Agradeço aos meus familiares, amigos, colegas, professores, orientadores e a todas as pessoas que me apoiaram, incentivaram e compartilharam sua sabedoria ao longo desse percurso. Cada gesto de apoio, cada palavra de encorajamento e cada momento de colaboração foram fundamentais para o sucesso deste trabalho.

Que esta conquista seja celebrada não apenas como um marco pessoal, mas também como uma vitória compartilhada por todos aqueles que estiveram ao meu lado. Que cada contribuição, por menor que pareça, seja reconhecida e valorizada como parte integrante deste processo. Que o aprendizado e as experiências vividas durante esta jornada continuem a inspirar e guiar meu caminho em novos desafios e conquistas que estão por vir.

A todos vocês, meu mais profundo agradecimento!

“Que todos os nossos esforços estejam sempre focados no desafio à impossibilidade. Todas as grandes conquistas humanas vieram daquilo que parecia impossível.”

(Charles Chaplin)

RESUMO

A Educação desempenha um papel fundamental no desenvolvimento das habilidades necessárias para enfrentar os desafios da vida contemporânea. Com o avanço da tecnologia, os recursos digitais têm ganhado cada vez mais relevância no processo educacional, oferecendo novas oportunidades para aprimorar a aprendizagem dos alunos. Nesse contexto, *Transfer Learning*, uma técnica de Aprendizado de Máquina, emerge como uma ferramenta promissora. Essa abordagem permite que modelos de aprendizagem, previamente treinados para resolver problemas específicos, sejam aplicados em novos contextos, ampliando sua utilidade e eficácia. O presente estudo propõe uma meta-análise para investigar e compreender a aplicação de *Transfer Learning* na Educação. Para isso, foi desenvolvido e executado um protocolo de Revisão Sistemática da Literatura (RSL), focado na identificação e extração de dados relevantes dos estudos selecionados. A meta-análise visa enriquecer os resultados obtidos na RSL, empregando técnicas estatísticas para uma análise mais aprofundada e robusta. Essa abordagem permitirá uma compreensão mais abrangente do uso de *Transfer Learning* na Educação, suas potencialidades e desafios enfrentados na prática. Ao integrar métodos qualitativos e quantitativos, este estudo fornece *insights* valiosos que contribuem para o avanço do conhecimento nessa área. A análise dos dados coletados permitiu identificar padrões, tendências e lacunas na literatura existente, além de destacar as principais aplicações e benefícios de *Transfer Learning* na Educação. Os resultados deste trabalho fornecem orientações úteis para pesquisadores, educadores e profissionais interessados em explorar e aplicar essa técnica inovadora em contextos educacionais.

Palavras-chave: Aprendizagem por Transferência. Educação. Revisão Sistemática da Literatura. Meta-análise.

ABSTRACT

Education plays a fundamental role in developing the skills necessary to face the challenges of contemporary life. With the advancement of technology, digital resources have gained increasing relevance in the educational process, offering new opportunities to enhance student learning. In this context, Transfer Learning, a Machine Learning technique, emerges as a promising tool. This approach allows previously trained learning models to be applied in new contexts, expanding their utility and effectiveness. The present study proposes a meta-analysis to investigate and understand the application of Transfer Learning in Education. For this purpose, a protocol for Systematic Literature Review (SLR) was developed and executed, focusing on the identification and extraction of relevant data from selected studies. The meta-analysis aims to enrich the results obtained in the SLR by employing statistical techniques for a more thorough and robust analysis. This approach will allow for a comprehensive understanding of the use of Transfer Learning in Education, its potentialities, and challenges faced in practice. By integrating qualitative and quantitative methods, this study seeks to provide valuable insights that contribute to advancing knowledge in this area. The analysis of collected data will identify patterns, trends, and gaps in existing literature, as well as highlight the main applications and benefits of Transfer Learning in Education. The results of this work provide useful guidance for researchers, educators, and professionals interested in exploring and applying this innovative technique in educational contexts.

Keywords: Transfer Learning. Education. Systematic Literature Review. Meta-analysis.

LISTA DE FIGURAS

Figura 1 – Teorias usadas na pesquisa.....	23
Figura 2 – Etapas da metodologia da pesquisa.....	33
Figura 3 – Visão geral do processo de seleção e extração de dados.	49
Figura 4 – Análise das palavras-chave extraídas dos 31 artigos selecionados.....	51
Figura 5 – Distribuição dos Estudos por Nacionalidade do Primeiro Autor	52
Figura 6 – Frequência dos 31 estudos da RSL por Etapa de Ensino.	64
Figura 7 – Visão geral da pontuação da avaliação de qualidade dos 31 artigos.....	86
Figura 8 – Distribuição dos Artigos por Categoria de Qualidade.....	89
Figura 9 – Mapa Mental sobre o estudo sobre Meta-análise.	91
Figura 10 – Análise SWOT da escolha do Método Aleatório em uma Meta-análise sobre <i>transfer learning</i> na Educação.	96
Figura 11 – Gráfico de Comparação de Desempenho entre Grupos de Estudos.	99
Figura 12 – Fluxo para Criação de um Gráfico de Floresta.....	105
Figura 13 – Gráfico de Floresta da Análise da Heterogeneidade nos Resultados de Estudos de <i>Transfer Learning</i> na Educação.....	107
Figura 14 – Fluxo para Criação da Análise de Variância (ANOVA).	111
Figura 15 – Fluxo para Criação do Teste de Tukey.....	117
Figura 16 – Fluxo para Criação do Teste de Levene.	121

LISTA DE TABELAS

Tabela 1 – Critérios de Inclusão e Exclusão.....	38
Tabela 2 – Listagem dos Dados para Extração dos Estudos Seleccionados na RSL. 40	
Tabela 3 – Listagem de Questões para Avaliação da Qualidade dos Estudos na RSL. 41	
Tabela 4 – Idealização das Hipóteses.	46
Tabela 5 – Síntese da Execução da <i>String</i> de Busca nas Fontes de Buscas.	48
Tabela 6 – Pontuação para cada nível de qualidade.....	87
Tabela 7 – Vantagens e Desvantagens do uso do <i>R</i> para conduzir Meta-análises	93
Tabela 8 – Comparação de Desempenho entre Grupos de Estudos.....	100
Tabela 9 – Estudos com Intervalo de Confiança.....	102
Tabela 10 – Resultado do Teste de Hipótese ANOVA.....	112
Tabela 11 – Resultado do Teste de Tukey	118
Tabela 12 – Resultado do Teste de Levene.....	122
Tabela 13 – Análise de variância (ANOVA) para grupos.	127
Tabela 14 – Resultado do Teste de Tukey	131

LISTA DE ABREVIATURAS E SIGLAS

ACM	<i>Association for Computing Machinery</i>
CNNs	Redes Neurais Convolucionais
DSB	<i>Daily School Break dataset</i>
ERNIE	<i>Enhanced Representation through kNowledge IntEgration</i>
IA	Inteligência Artificial
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
MAS	<i>Multiagent System</i>
NDCG	<i>Normalized Discounted Cumulative Gain</i>
QP	<i>Questões de Pesquisa</i>
RL	<i>Reinforcement Learning</i>
RSL	Revisão Sistemática da Literatura
SWOT	<i>Strengths, Weaknesses, Opportunities, Threats</i>
UFSC	Universidade Federal de Santa Catarina

SUMÁRIO

1	INTRODUÇÃO	18
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	Educação	23
2.2	<i>Transfer Learning</i>	24
2.3	Revisão Sistemática	25
2.4	Meta-análise	26
2.5	Considerações Finais do Capítulo	27
3	TRABALHOS RELACIONADOS	29
3.1	Considerações Finais do Capítulo	32
4	METODOLOGIA	33
4.1	Revisão Sistemática da Literatura	34
4.2	Planejamento do Protocolo RSL	34
4.3	Definição do Objetivo e Questões de Pesquisa	35
4.3.1	<i>Objetivo</i>	35
4.3.2	<i>Questões de Pesquisa</i>	35
4.4	Elaboração da <i>String</i> de Busca e definição de Critérios de Seleção .	36
4.4.1	<i>String de Busca</i>	37
4.4.2	<i>Critérios de Seleção</i>	37
4.5	Método de Pesquisa	38
4.6	Extração e Sumarização dos Dados, e Discussão dos Resultados	39
4.7	Meta-análise	41
4.7.1	<i>Estudo sobre Meta-análise</i>	42
4.7.2	<i>Identificação de Softwares</i>	43
4.7.3	<i>Definição do Modelo Estatístico</i>	44
4.7.4	<i>Avaliação da Heterogeneidade</i>	44
4.7.5	<i>Idealização de Testes de Hipóteses</i>	45
4.8	Identificação de Lacunas	46
4.9	Considerações Finais do Capítulo	47
5	RESULTADOS	48
5.1	Visão Geral dos Resultados da RSL	48

5.2	Discussões Gerais	51
5.2.1	<i>Locais de Publicação e Nacionalidade do Primeiro Autor</i>	51
5.2.1.1	<i>Objetivos e Metodologias</i>	55
5.2.1.2	<i>Métodos e Técnicas</i>	58
5.2.1.3	<i>Contexto Educacional</i>	62
5.2.1.4	<i>Conjuntos de Dados</i>	65
5.2.1.5	<i>Aplicação de Transfer Learning</i>	69
5.2.1.6	<i>Desempenho dos Modelos</i>	73
5.2.1.7	<i>Principais Resultados e Descobertas</i>	77
5.2.1.8	<i>Conclusões e Validações</i>	80
5.2.1.9	<i>Lacunas de Pesquisa</i>	82
5.3	Avaliação de Qualidade dos Artigos Seleccionados na RSL	84
5.4	Visão Geral dos Resultados da Meta-análise.....	90
5.4.1	<i>Resultado do Estudo sobre Meta-análise</i>	90
5.4.2	<i>Resultado da Identificação de Softwares</i>	91
5.4.3	<i>Resultado da Definição de Modelos</i>	94
5.4.4	<i>Resultado da Avaliação da Heterogeneidade dos Estudos</i>	97
5.4.4.1	<i>Comparação de Desempenho entre Grupos</i>	97
5.4.4.2	<i>Análise da Heterogeneidade com Gráfico de Floresta</i>	103
5.4.5	<i>Resultado dos Testes de Hipóteses em Meta-Análises</i>	109
5.4.5.1	<i>Cenário 1: Divisão de Grupos baseada em Métricas de Desempenho .</i>	109
5.4.5.2	<i>Cenário 2: Divisão dos Grupos baseada em Níveis Educacionais</i>	123
5.4.5.3	<i>Discussão dos Resultados dos Testes de Hipóteses nos Cenários 1 e 2 .</i>	133
6	CONSIDERAÇÕES FINAIS.....	135
	REFERÊNCIAS	141
	APÊNDICES.....	149
APÊNDICE A	ARTIGOS SELECIONADOS NA RSL	150
APÊNDICE B	ARTIGOS INDISPONÍVEIS NA RSL	153
APÊNDICE C	DADOS EXTRAÍDOS DOS ARTIGOS	154
APÊNDICE D	DATASETS UTILIZADOS NOS ARTIGOS.....	173
APÊNDICE E	DADOS (OUTROS) EXTRAÍDOS DOS ARTIGOS	184

1 INTRODUÇÃO

No panorama educacional atual, é notável o constante fluxo de inovações, que o moldam em um cenário dinâmico e em constante evolução. Segundo Sousa e Moura (2018), a rápida progressão da tecnologia tem desempenhado um papel significativo na transformação da maneira como as informações são acessadas, compartilhadas e processadas. A cada momento, novos avanços tecnológicos são previstos, impulsionados pela crescente disponibilidade e uso desses recursos, com o objetivo primordial de enriquecer a experiência de aprendizagem dos alunos. Silva (2020) afirma que métodos educacionais são aplicados estrategicamente para otimizar o processo de ensino e aprendizagem, enquanto Gomes *et al.* (2019) enfatiza a abrangência dessas estratégias para alcançar melhores resultados educacionais.

À medida que o mundo se torna cada vez mais interconectado, a globalização impõe uma demanda crucial aos sistemas educacionais: preparar os estudantes para operarem efetivamente em um ambiente global em constante evolução. Conforme observado por Castells (2000), a educação deve evoluir para fornecer aos alunos as habilidades necessárias para navegar e prosperar na sociedade em rede emergente, onde a informação e a comunicação desempenham papéis centrais. Nesse cenário dinâmico, surge a necessidade de sistemas educacionais inovadores, capazes de fornecer suporte personalizado e adaptativo.

A integração da Inteligência Artificial (IA) nesse contexto desempenha um papel fundamental, permitindo uma compreensão mais profunda das emoções dos alunos durante o processo de aprendizado. Conforme observado por Chen *et al.* (2020), a aplicação da IA na educação está transformando a maneira como é entendido e apoiado os processos de aprendizado, tornando-os mais dinâmicos e responsivos às necessidades dos estudantes. A capacidade de identificar padrões que indiquem possíveis dificuldades de aprendizado e implementar intervenções específicas em tempo real torna-se, portanto, uma ferramenta essencial para aprimorar a abordagem educacional e promover o sucesso acadêmico de todos os estudantes.

As inovações tecnológicas estão abrindo novas perspectivas para métodos de ensino mais dinâmicos e interativos. Com os avanços na área de IA, surge a oportunidade de explorar novas abordagens no contexto educacional. Um exemplo promissor é a aprendizagem por transferência, que tem sido objeto de diversos estudos recentes. Essa abordagem envolve a reutilização do conhecimento adquirido ao resolver um determinado problema

para aprimorar a resolução de uma nova tarefa. Conforme discutido por Torrey e Shavlik (2010), a Aprendizagem por Transferência (do inglês, *Transfer Learning*) permite melhorar a performance em uma nova tarefa ao aplicar conhecimentos previamente adquiridos. Essencialmente, essa metodologia implica o aproveitamento de modelos pré-treinados como ponto de partida para resolver desafios distintos.

A Aprendizagem por Transferência representa uma abordagem fascinante no campo da inteligência artificial e do aprendizado de máquina. Pan e Yang (2010) destacam que essa abordagem tem o potencial de revolucionar a forma como aborda-se problemas de aprendizado de máquina, tornando os modelos mais versáteis e adaptáveis. Essa técnica se baseia na ideia de que conhecimentos adquiridos ao resolver um problema específico podem ser transferidos e aplicados para melhorar o desempenho em tarefas relacionadas. Por exemplo, um modelo treinado para reconhecer rostos em imagens pode ser adaptado com sucesso para reconhecer expressões faciais em vídeos.

Na Educação, a aprendizagem por transferência oferece promessas significativas, pois permite que os alunos aproveitem o que aprenderam em um contexto para melhorar seu desempenho em situações semelhantes. Weiss *et al.* (2016) destacam que a aprendizagem por transferência não apenas aprimora o desempenho acadêmico, mas também prepara os alunos para enfrentar problemas complexos de maneira inovadora e adaptável. Isso pode ser especialmente útil em disciplinas como matemática e ciências, onde conceitos fundamentais podem ser aplicados em uma variedade de problemas. No entanto, apesar de seu potencial, ainda há desafios a serem enfrentados, como a seleção adequada de modelos pré-treinados e a adaptação eficaz para diferentes contextos educacionais. À medida que a exploração e o refinamento dessa técnica continuam, esperam-se avanços significativos no campo da educação, possibilitando uma aprendizagem mais eficaz e personalizada para os alunos.

A educação se adapta aos desafios impostos pela era digital e pela globalização, cresce a necessidade por abordagens inovadoras que otimizem o processo de ensino e aprendizagem. Nesse contexto, o uso do aprendizado por transferência na educação desponta como uma área de grande interesse, prometendo aproveitar conhecimentos prévios para aprimorar a eficácia do ensino. Diante desse cenário desafiador, torna-se essencial realizar uma revisão sistemática e uma meta-análise sobre o tema. Esse empreendimento não apenas consolidaria o conhecimento existente sobre as práticas de

transfer learning na educação, mas também permitirá identificar lacunas de pesquisa e áreas de oportunidade para estudos futuros. Ao sintetizar e analisar criticamente os resultados de estudos anteriores, essa revisão sistemática e meta-análise poderiam fornecer *insights* valiosos para educadores, pesquisadores e formuladores de políticas, contribuindo assim para o contínuo avanço do campo educacional em direção a práticas mais eficazes e adaptativas.

Enquanto a revisão sistemática da literatura fornece uma compilação abrangente dos estudos existentes sobre *transfer learning* na educação, a realização de uma meta-análise agrega um novo nível de *insight* ao sintetizar quantitativamente os resultados desses estudos. Nesse contexto, o protocolo de estudo surge como um processo para orientar a condução de uma pesquisa, planejando e documentando todas as etapas para análise e obtenção de resultados. Brizola e Fantin (2016) relatam em sua pesquisa que a Revisão Sistemática possibilita o levantamento de diversas informações e, conseqüentemente, gera discussões sobre vários aspectos abordados nos estudos.

Nesta etapa complementar de investigação, será conduzido uma meta-análise para quantificar e avaliar o impacto dos modelos de aprendizado por transferência em comparação com os métodos tradicionais de ensino e aprendizagem em diversos contextos educacionais. Embora a revisão sistemática ofereça uma visão panorâmica das descobertas na literatura, a meta-análise permite ir além da mera descrição qualitativa, fornecendo uma avaliação quantitativa e estatisticamente robusta do efeito do *transfer learning* na educação. Ao combinar os resultados de estudos individuais, a meta-análise nos capacita a identificar padrões, tendências e diferenças significativas entre os grupos de estudo, ampliando assim nossa compreensão sobre a eficácia relativa dessas abordagens no contexto educacional.

A problemática abordada nesta revisão sistemática da literatura e meta-análise reside na necessidade de compreender o impacto e a eficácia do aprendizado por transferência na educação. Embora o conceito de *transfer learning* tenha ganhado destaque como uma estratégia promissora para aprimorar o processo de ensino e aprendizagem, há uma lacuna significativa no entendimento de suas aplicações específicas, suas limitações e seu potencial para diferentes contextos educacionais. Além disso, a diversidade de abordagens metodológicas e resultados inconsistentes em estudos individuais podem dificultar a obtenção de uma visão clara e abrangente sobre o tema. Portanto, a problemática central é identificar as tendências, padrões e descobertas comuns emergentes dos estudos revisados,

a fim de oferecer orientações valiosas para educadores, pesquisadores e formuladores de políticas educacionais sobre como melhor utilizar o aprendizado por transferência para aprimorar a prática educacional e promover resultados de aprendizagem mais eficazes.

O objetivo desta revisão sistemática da literatura e meta-análise é proporcionar uma visão abrangente dos resultados dos estudos que investigaram a aplicação do aprendizado por transferência na educação. O intuito é identificar tendências, padrões e descobertas comuns emergentes dos estudos revisados, visando aprofundar a compreensão sobre o papel e o impacto do *transfer learning* no contexto educacional. Foi adotado uma abordagem meticulosa e sistemática para identificar estudos relevantes relacionados ao tema. Foram realizadas buscas em bases de dados eletrônicas utilizando termos específicos relacionados ao aprendizado por transferência e educação. Os critérios de inclusão e exclusão foram aplicados de forma criteriosa para selecionar os estudos mais pertinentes e de alta qualidade. Adicionalmente, foi conduzido uma análise crítica dos estudos selecionados, com o objetivo de extrair e sintetizar os principais resultados de maneira objetiva e rigorosa. Essa abordagem permitirá oferecer uma visão aprofundada e fundamentada sobre o estado atual da pesquisa nesse campo em constante evolução.

Neste trabalho, será apresentada uma Revisão Sistemática da Literatura sobre o *Transfer Learning* na educação, destacando seu uso para melhorar o desempenho em novas tarefas. Embora essa abordagem demonstre potencial para aprimorar a eficácia do ensino, enfrenta desafios e questões, incluindo a heterogeneidade nos estudos sobre o assunto. Essa heterogeneidade decorre das diferenças nos contextos educacionais, na diversidade das tarefas educacionais, nas modalidades de aprendizado e nas técnicas de *Transfer Learning* empregadas. A partir desta pesquisa, busca-se enriquecer os resultados obtidos na RSL através do uso da meta-análise, empregando testes estatísticos. Dessa forma, esperamos contribuir positivamente incentivando os pesquisadores a abordarem as lacunas identificadas, propondo novos estudos que envolvam o *Transfer Learning* e a educação. Almejamos superar esses desafios, padronizando métodos, relatando claramente os contextos de estudo, considerando a diversidade de tarefas e populações para facilitar a reprodutibilidade e a comparação entre estudos.

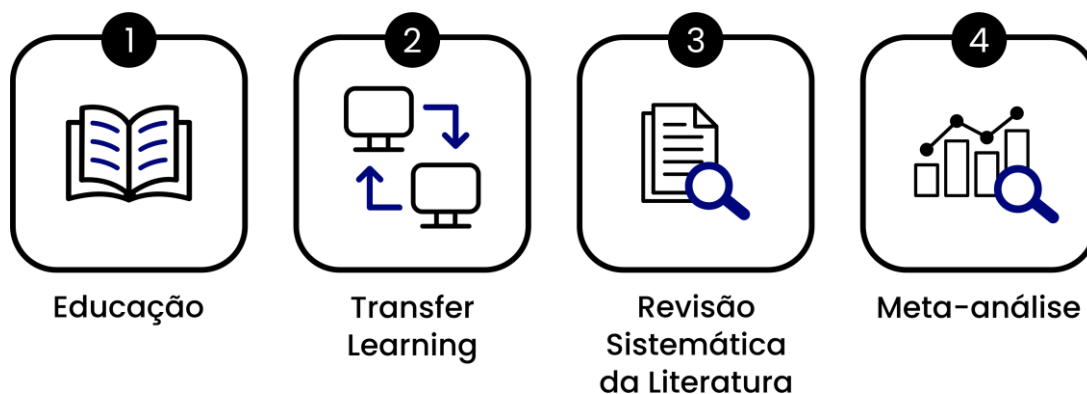
O trabalho segue uma estrutura organizada para facilitar a compreensão e o acompanhamento dos temas abordados. No Capítulo 2, é apresentada a fundamentação teórica, fornecendo os conceitos e princípios essenciais para a compreensão do tema em questão. O

Capítulo 3 descreve os trabalhos relacionados, oferecendo uma visão abrangente do estado atual da pesquisa na área e destacando contribuições relevantes de estudos anteriores. O Capítulo 4 detalha os procedimentos metodológicos adotados, delineando o caminho seguido para a realização do estudo. No Capítulo 5, são apresentados os resultados da Revisão Sistemática da Literatura e da Meta-análise, fornecendo insights que orientam a discussão e análises. O Capítulo 6 contém a conclusão e aponta direções para trabalhos futuros, consolidando os achados do estudo e delineando possíveis caminhos para pesquisas subsequentes. Além disso, são incluídas as referências bibliográficas, fundamentais para embasar as informações apresentadas, e os apêndices A ao E, que oferecem espaço adicional para o detalhamento de informações relevantes para uma compreensão completa do trabalho. Essa estrutura foi cuidadosamente elaborada para proporcionar uma leitura fluida e compreensível, permitindo uma imersão completa nos temas abordados.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, serão abordados os conceitos e teorias fundamentais para uma compreensão aprofundada e a realização eficaz deste trabalho (ver Figura 1).

Figura 1 – Teorias usadas na pesquisa.



Fonte: Autoria Própria (2024).

2.1 Educação

A educação é essencialmente o processo de transmitir conhecimento e instrução. Conforme estabelecido pela Lei nº 9.394, de 20 de dezembro de 1996, que define as Diretrizes e Bases da Educação Nacional, a educação compreende os procedimentos formativos desenvolvidos ao longo da vida, principalmente por meio do ensino em instituições educacionais dedicadas. Vasconcelos (2007) destaca o papel crucial da escola na formação cidadã. Segundo o autor, a escola não apenas transmite conhecimento acadêmico, mas também desempenha um papel fundamental na promoção de valores, habilidades sociais e cívicas essenciais para o desenvolvimento integral dos alunos.

O termo "educação" abrange o ato de educar, ou seja, instruir indivíduos para a aquisição de conhecimento. Trata-se de um processo contínuo de aprendizado e desenvolvimento que visa a obtenção de conhecimentos, habilidades, valores, atitudes e competências. Rocha (2016) ressalta que a educação é um instrumento fundamental para o desenvolvimento pessoal e coletivo, contribuindo para a compreensão do mundo, a promoção da igualdade, o crescimento econômico e o avanço cultural. Para que a educação seja eficaz, é essencial contar com um ambiente escolar apropriado, seja público ou privado, destinado a fornecer educação coletiva. Esteban (2007) observa que, no Brasil, as escolas

públicas predominam entre crianças, jovens e adultos das classes populares.

A educação enfrenta diversos desafios históricos, buscando constantemente melhorar a qualidade do ensino. Um marco significativo foi a implementação do ensino integral, com a ampliação das atividades escolares. Cavaliere (2007) destaca o aumento dos projetos de educação integral na rede pública brasileira, caracterizada pela adoção do ensino em período integral.

Neste trabalho, será explorado a aprendizagem por transferência, uma variante da aprendizagem supervisionada aplicada à educação, especialmente em contextos relacionados à qualidade do ensino-aprendizagem. O objetivo é investigar o desempenho dessa técnica aplicada a contextos educacionais, identificando e analisando modelos de aprendizado de máquina adaptados para resolver problemas dentro desse contexto. Esses modelos podem ser reutilizados como soluções base para novos problemas dentro do mesmo contexto ou em contextos semelhantes.

2.2 *Transfer Learning*

O conceito da técnica de Aprendizagem por Transferência (ou *Transfer Learning*, em inglês) é intrinsecamente ligado ao armazenamento de conhecimento adquirido na resolução de problemas existentes, visando sua aplicação em contextos semelhantes, mas distintos. Durante seu uso, observa-se que, por meio do treinamento, o modelo é capaz de adquirir novas habilidades e utilizar modelos preexistentes para abordar novos problemas. Ying *et al.* (2018) afirmam que essa técnica visa potencializar o aprendizado, utilizando um modelo original para aprimorar o desempenho no aprendizado do domínio de destino, transferindo o conhecimento entre tarefas.

De acordo com Zhuang *et al.* (2020), a transferência de aprendizado visa melhorar o desempenho do conhecimento adquirido, transferindo-o entre tarefas para potencializar o aprendizado. Assim, a aplicação refere-se à habilidade de aplicar o conhecimento adquirido em experiências anteriores a novas situações, influenciando seu desempenho em um novo contexto, aproveitando as semelhanças entre componentes, padrões e contextos.

Para aplicar o modelo pré-treinado de *Transfer Learning* a um novo problema, é crucial realizar o pré-processamento dos dados, preparando-os para o treinamento da rede neural. Em seguida, a leitura e o tratamento dos dados são realizados, eliminando arquivos incorretos, se houver. Além disso, é essencial definir os rótulos (*labels*) para o treinamento

e, assim, treinar as Redes Neurais Convolucionais (CNNs).

Segundo Pan e Yang (2009), a transferência de conhecimento pode trazer benefícios significativos ao processo de aprendizado, permitindo um melhor desenvolvimento e desempenho sem a necessidade de custosos esforços de rotulagem de dados. No entanto, sua aplicabilidade é complexa, dada a dificuldade de aplicá-la em situações novas e diferentes, exigindo uma consideração cuidadosa dos fatores que podem influenciar sua eficácia.

Em suma, a Aprendizagem por Transferência emerge como uma técnica promissora para aprimorar o aprendizado em diferentes contextos educacionais, permitindo a transferência de conhecimento adquirido de tarefas anteriores para novos problemas. A aplicação dessa técnica requer cuidadoso pré-processamento de dados e treinamento de redes neurais, mas oferece benefícios significativos, como a melhoria do desempenho do aprendizado sem a necessidade de rotulagem extensiva de dados. No entanto, sua complexidade de aplicação exige uma consideração minuciosa dos fatores que podem influenciar sua eficácia.

2.3 Revisão Sistemática

A Revisão Sistemática da Literatura (RSL) desempenha um papel fundamental na condução de pesquisas científicas, oferecendo uma maneira estruturada de compreender o panorama de estudos relacionados a um tema específico e facilitando a discussão sobre modelos, métodos e técnicas utilizados em pesquisas anteriores. Como destacado por Dermeval *et al.* (2020), essa revisão visa reduzir tanto os erros sistemáticos quanto os aleatórios, e sua realização requer a aplicação de estratégias que permitam que sua abrangência seja avaliada por outros pesquisadores.

Além de fornecer uma visão abrangente do conhecimento existente, a RSL também serve como uma ferramenta para identificar lacunas nas pesquisas prévias, contribuindo para o desenvolvimento de novos estudos. Como mencionado por Donato e Donato (2019), a RSL é conduzida por meio de métodos sistemáticos que envolvem a formulação de perguntas de pesquisa, busca em bibliotecas digitais, seleção de artigos relevantes, extração de dados e avaliação da qualidade metodológica das pesquisas selecionadas.

Para facilitar o processo de busca por pesquisas, é importante organizar um protocolo de revisão, para definir como esse processo será conduzido. É importante seguir

alguns passos, como:

1. Formulação de perguntas de pesquisa: Definir perguntas claras e específicas que orientarão a busca e a análise dos estudos relacionados ao tema de interesse.
2. Busca nas bibliotecas digitais da literatura: Realizar uma busca sistemática em bases de dados científicas, como PubMed, Scopus ou Web of Science, utilizando termos relevantes ao tema da pesquisa.
3. Seleção dos artigos: Aplicar critérios de inclusão e exclusão para selecionar os artigos que atendam aos objetivos da revisão e à sua pertinência ao tema em questão.
4. Extração dos dados: Extrair informações relevantes de cada artigo selecionado, de acordo com as perguntas de pesquisa definidas anteriormente.
5. Avaliação da qualidade metodológica: Realizar uma avaliação crítica da qualidade metodológica dos estudos incluídos na revisão, identificando possíveis vieses e limitações.

Ao seguir esses passos de maneira sistemática e rigorosa, é possível conduzir uma revisão abrangente e confiável da literatura, fornecendo uma base sólida para a análise e discussão dos resultados. O protocolo de revisão, conforme destacado por Donato e Donato (2019), é essencial para orientar o processo e garantir a eficiência da revisão. Portanto, a adoção de um protocolo de revisão sistemática da literatura é fundamental para a condução de pesquisas de alta qualidade, fornecendo uma base sólida para a tomada de decisões e o avanço do conhecimento científico em uma determinada área.

2.4 Meta-análise

A meta-análise, uma técnica estatística fundamental, complementa a revisão sistemática da literatura (RSL) ao agregar e analisar quantitativa e qualitativamente os resultados dos estudos revisados. Como mencionado por Fuchs e Paim (2010), ela é essencialmente uma revisão sistemática, porém, incorpora métodos estatísticos para oferecer uma visão mais ampla e detalhada dos resultados obtidos.

Amplamente aplicada em diversas áreas do conhecimento, a meta-análise integra os resultados dos estudos, fornecendo um resumo estatístico que facilita a compreensão e a interpretação das descobertas. Conforme destacado por Munoz *et al.* (2002), essa técnica é um instrumento metodológico importante, pois permite uma análise numérica abrangente, abordando múltiplos trabalhos de pesquisa e impulsionando o avanço do conhecimento de

forma eficaz e contínua.

Através da meta-análise, é possível realizar análises detalhadas dos dados, explorando as diferenças metodológicas entre os estudos revisados e propondo padrões que justifiquem e expliquem essas divergências. Conforme evidenciado por Gonçalves *et al.* (2015), essa técnica oferece uma compreensão mais abrangente das questões em estudo, ao mesmo tempo em que ressalta a importância dos níveis de evidência e validade.

No processo de meta-análise, a escolha do modelo estatístico adequado é crucial. Geralmente, são empregados dois tipos de modelos: o de efeito fixo e o de efeito aleatório. Como mencionado por Fuchs e Paim (2010), essa escolha é determinante para a análise dos dados, uma vez que as diferenças entre os estudos podem ser exploradas estatisticamente para gerar discussões mais robustas.

O modelo de efeito fixo assume que todos os estudos compartilham um efeito comum, ou seja, qualquer variação observada entre eles é atribuída ao acaso. Já o modelo de efeito aleatório reconhece que os estudos podem ter efeitos diferentes, considerando que essa variação é real e não apenas fruto de erros amostrais. Como apontado por Marques *et al.* (2000), essas diferenças influenciam tanto a interpretação dos resultados quanto a escolha da abordagem estatística.

Além disso, é importante considerar o conceito de heterogeneidade, que reflete as diferenças metodológicas entre os estudos e pode indicar a presença de viés. Conforme observado por Santos e Cunha (2016), a heterogeneidade permite uma avaliação mais precisa da variabilidade dos resultados e das possíveis influências sobre os efeitos observados.

Por fim, a aplicação de testes de hipóteses é fundamental para validar as afirmações e interpretar os dados de forma objetiva. Conforme explicado por Hirakata *et al.* (2019), os testes de hipóteses são ferramentas essenciais na análise estatística, fornecendo suporte para a tomada de decisões e a interpretação dos resultados obtidos na pesquisa.

2.5 Considerações Finais do Capítulo

Considerando o conteúdo abordado no capítulo de Fundamentação Teórica, torna-se evidente a importância da integração da informática na educação, não apenas como uma ferramenta, mas como um facilitador essencial para o engajamento dos estudantes e o desenvolvimento de competências multifacetadas. A interseção entre a informática e a educação promove um ambiente de aprendizado dinâmico, onde os alunos podem explorar,

criar e colaborar, preparando-se para os desafios do mundo moderno.

Nesse contexto, o conceito de *Transfer Learning* surge como uma abordagem poderosa, capaz de reutilizar conhecimentos adquiridos em tarefas relacionadas para melhorar o desempenho em novos contextos. Essa capacidade de transferência de aprendizado não só impulsiona a eficácia do ensino, mas também promove uma abordagem mais adaptativa e personalizada para atender às necessidades individuais dos alunos.

Além disso, a Revisão Sistemática da Literatura (RSL) emerge como uma ferramenta essencial no arsenal do pesquisador, permitindo a síntese e análise de uma vasta gama de estudos relacionados a um determinado tema. Através da RSL, é possível mapear o conhecimento existente, identificar lacunas de pesquisa e estabelecer uma base sólida para estudos futuros.

Por fim, a meta-análise oferece uma abordagem estatística robusta para integrar e analisar os resultados de múltiplos estudos, proporcionando uma visão mais abrangente sobre o tema em questão. Ao combinar dados quantitativos e qualitativos, a meta-análise capacita os pesquisadores a extrair conclusões significativas e fundamentadas, impulsionando o avanço do conhecimento em diversas áreas.

Portanto, é evidente que a informática na educação, juntamente com abordagens como *Transfer Learning*, Revisão Sistemática da Literatura e meta-análise, desempenham papéis essenciais na promoção da inovação, excelência e avanço contínuo no campo educacional. Ao integrar essas ferramentas e conceitos de forma estratégica, pode-se impulsionar o desenvolvimento de soluções educacionais mais eficazes, adaptativas e centradas no aluno.

3 TRABALHOS RELACIONADOS

Neste capítulo são apresentados e discutidos os trabalhos relacionados ao contexto desta pesquisa, explorando processos e metodologias para a realização de revisões sistemáticas da literatura e meta-análises.

Os trabalhos relacionados apresentados a seguir foram coletados em diversas fontes confiáveis de pesquisa acadêmica, como a Biblioteca Digital da *Association for Computing Machinery* (ACM), que abriga uma vasta coleção de documentos nas áreas de Computação e Tecnologia da Informação; a Biblioteca Digital do *Institute of Electrical and Electronics Engineers* (IEEE), uma fonte importante para acessar conteúdos científicos e técnicos; e a plataforma Scopus, conhecida por indexar artigos científicos publicados em periódicos acadêmicos. Essas fontes fornecem uma base sólida para a identificação e análise de trabalhos relevantes no contexto desta pesquisa.

Diversos estudos têm se dedicado a identificar e analisar soluções computacionais que reutilizam modelos preditivos ajustados para problemas específicos, visando identificar o estado da arte de pesquisas relevantes sobre o uso de aprendizagem por transferência. Tais estudos abordam questões cruciais, como os principais problemas enfrentados nas pesquisas, a aplicabilidade do *transfer learning* no contexto educacional e os métodos de validação das soluções propostas.

Um mapeamento sistemático e uma revisão sistemática da literatura em informática na educação foram conduzidos por Dermeval *et al.* (2020), estabelecendo diretrizes fundamentais para a realização dessas revisões. Por meio de um protocolo meticulosamente definido, o estudo abarcou etapas que vão desde o planejamento até a redação do relatório final, proporcionando um guia abrangente e detalhado para pesquisadores que buscam realizar pesquisas nessa área. Este trabalho se destaca pela sua organização rigorosa e pela ênfase na metodologia, contribuindo significativamente para a padronização e a qualidade das revisões na área de informática educativa.

O estudo conduzido por Passos *et al.* (2006) representa uma contribuição significativa para compreender o desenvolvimento profissional dos professores de matemática no contexto brasileiro. Por meio de uma abordagem metódica, o trabalho realizou uma meta-análise de uma ampla gama de estudos acadêmicos, incluindo dissertações de mestrado e teses de doutorado, com o propósito de investigar práticas que promovem o desenvolvimento profissional dos educadores. Ao integrar essas pesquisas, o estudo

gerou discussões essenciais sobre a formação e a evolução profissional dos professores, destacando as práticas reflexivas, investigativas e colaborativas como elementos-chave nesse processo. A análise minuciosa conduzida permitiu identificar evidências sólidas que reforçam a importância dessas práticas na promoção do desenvolvimento profissional dos professores de matemática, fornecendo insights valiosos para aprimorar a educação nessa área.

O estudo realizado por SPANHOL *et al.* (2010) representa uma investigação abrangente sobre o estado atual da educação a distância, visando identificar suas contribuições significativas. Utilizando o banco de dados da Biblioteca Universitária (BU) da Universidade Federal de Santa Catarina (Universidade Federal de Santa Catarina (UFSC)), que reúne todas as teses e dissertações defendidas na instituição, os pesquisadores conduziram uma busca criteriosa por trabalhos disponíveis entre os anos de 1998 e 2008, utilizando termos relevantes como "educação a distância", "ensino a distância" ou "EaD". Os dados coletados foram submetidos a uma análise estatística detalhada, incluindo medidas como frequência absoluta, frequência relativa e média. Além disso, foram realizados cruzamentos de dados para análises mais abrangentes. Essa metodologia permitiu uma avaliação precisa da produção acadêmica relacionada à educação a distância na instituição, oferecendo insights valiosos sobre a colaboração e as tendências nessa área específica de pesquisa.

O estudo conduzido por Guimarães *et al.* (2020) representa uma revisão sistemática da literatura sobre a produção de *Learning Analytics* e a Predição de Desempenho Acadêmico por pesquisadores brasileiros, utilizando a ferramenta Parsifal. O principal objetivo dessa pesquisa foi identificar, avaliar e interpretar toda a pesquisa disponível nessa área. Para alcançar esse propósito, os pesquisadores empregaram uma metodologia robusta, elaborando um protocolo de extração de dados específico para os resultados de predição de desempenho de aprendizagem, além de identificar perfis e realizar testes utilizando algoritmos de predição. A abrangência da revisão incluiu estudos publicados em diversas bases de pesquisa renomadas, como Scopus, Science Direct, Scielo, IEEE e CEIE. Através desse processo meticuloso, foram selecionados diversos artigos para compor a pesquisa, permitindo uma análise aprofundada das técnicas e algoritmos utilizados na predição de desempenho acadêmico.

O estudo conduzido por Taylor e Stone (2009) representa uma pesquisa do tipo

survey, caracterizada por uma investigação qualitativa que teve como objetivo principal abordar os desafios relacionados à aprendizagem por transferência em domínios de Aprendizado por Reforço (Reinforcement Learning - *Reinforcement Learning* (RL)). Essa pesquisa buscou identificar os métodos de transferência mais relevantes na área, organizar e discutir questões consideradas importantes nesse contexto. A pesquisa explorou a aprendizagem de Transferência para Domínios de Aprendizagem por Reforço, apresentando um framework que associa os métodos relacionados aos termos de suas capacidades e objetivos. Essa estrutura foi empregada para realizar uma análise detalhada da literatura existente, permitindo identificar lacunas de pesquisa e sugerir direções para trabalhos futuros.

O estudo conduzido por Silva e Costa (2019) abordou o aprendizado por transferência para sistemas de aprendizado por reforço multiagentes, com o objetivo de promover a utilização eficiente de todas as fontes de conhecimento disponíveis em Sistemas Multiagentes (Multiagent Systems - *Multiagent System* (MAS)). Essa pesquisa ofereceu uma visão unificada da literatura sobre o reuso de conhecimento em RL multiagente, fornecendo uma análise abrangente das soluções existentes para o problema geral de reutilização de conhecimento. Para isso, foi estabelecida uma taxonomia de soluções que abordava diversas técnicas e estratégias de reuso de conhecimento entre agentes, permitindo uma compreensão mais clara do progresso recente nessa área.

Diante dos estudos mencionados, fica evidente a necessidade de uma pesquisa abrangente que integre análises quantitativas e qualitativas sobre a aplicação da técnica de *transfer learning* na área da Educação. Até o momento, não foi identificado um conjunto significativo de estudos que aborde essa temática de forma abrangente. Assim, é fundamental conduzir uma investigação que vá além das análises fragmentadas presentes na literatura atual, buscando compreender tanto os aspectos quantitativos, como métricas de desempenho e resultados estatísticos, quanto os aspectos qualitativos, como impacto pedagógico, aceitação pelos usuários e considerações éticas.

Uma pesquisa que englobe esses dois domínios oferecerá uma visão mais completa e profunda sobre o potencial do *transfer learning* na Educação, permitindo uma avaliação mais abrangente de sua eficácia, desafios e oportunidades. Essa abordagem integrada é crucial para fornecer insights significativos e orientar o desenvolvimento futuro de aplicações de *transfer learning* voltadas para aprimorar os processos de ensino e aprendizagem.

3.1 Considerações Finais do Capítulo

Com base nos trabalhos relacionados apresentados neste capítulo, fica claro que correlacionar estudos anteriores com a presente pesquisa é crucial para enriquecer seu embasamento teórico e explorar novas conjecturas. Essa abordagem permite não apenas identificar lacunas existentes no conhecimento, mas também fornecer insights valiosos que podem influenciar significativamente o desenvolvimento do estudo em questão.

Ao analisar os diversos estudos que empregam soluções computacionais, especialmente nas revisões sistemáticas na área de informática na educação, é possível observar as principais etapas envolvidas em sua condução. Essas etapas, que abrangem desde o planejamento até a síntese dos resultados, destacam a importância de uma metodologia rigorosa para garantir a confiabilidade e a validade dos achados obtidos.

Além disso, a utilização da meta-análise como uma ferramenta para agregar e analisar dados estatísticos provenientes de múltiplos estudos demonstra a importância de se obter uma visão global e sumariada das evidências disponíveis. Esse enfoque estatístico permite extrair conclusões mais robustas e generalizáveis, contribuindo para o avanço do conhecimento em determinada área.

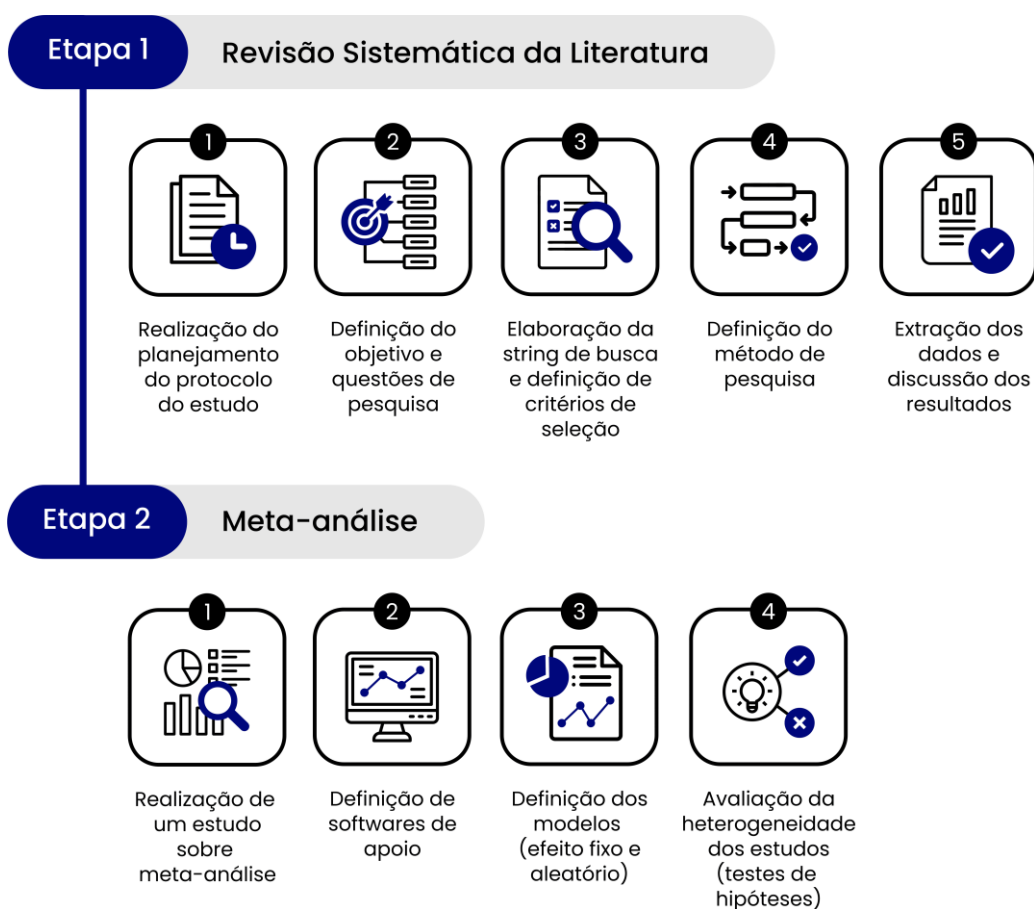
Outro ponto relevante observado nos trabalhos relacionados é a crescente ênfase atribuída ao *Transfer Learning* como uma estrutura que facilita a reutilização e a transferência de conhecimento entre diferentes contextos. Essa abordagem demonstra ser uma ferramenta poderosa para maximizar o aproveitamento de diversas fontes de informação e promover uma aprendizagem mais eficiente e adaptativa.

Portanto, ao considerar as contribuições dos estudos relacionados, torna-se evidente a relevância de adotar uma abordagem abrangente e criteriosa na condução da presente pesquisa, a fim de explorar plenamente o potencial do *Transfer Learning* e contribuir de forma significativa para o avanço do conhecimento na área da Educação e da Informática.

4 METODOLOGIA

Neste capítulo são apresentadas as atividades realizadas para o alcance dos objetivos deste trabalho. Para tanto, utilizou-se uma abordagem de análise da literatura em duas etapas distintas, conforme ilustrado na Figura 2. Neste caso, conduziu-se uma Revisão Sistemática da Literatura (RSL) com um protocolo de pesquisa bem definido, seguida por uma Meta-Análise que englobou um conjunto de medidas e testes estatísticos baseados nos estudos identificados na RSL.

Figura 2 – Etapas da metodologia da pesquisa.



Fonte: Autoria Própria (2024).

Essas etapas metodológicas forneceram uma base sólida para o desenvolvimento do presente estudo, possibilitando uma análise abrangente e fundamentada sobre o tema em questão. A seguir, detalha-se cada uma dessas etapas, com os procedimentos específicos a serem seguidos para suas execuções. Em resumo, a combinação de uma Revisão Sistemática da Literatura com uma meta-análise promove avanços significativos no conhecimento da área e oferece subsídios valiosos para decisões futuras.

4.1 Revisão Sistemática da Literatura

Neste estudo, a Revisão Sistemática da Literatura foi realizada em cinco etapas estruturadas: (1) planejamento do protocolo; (2) definição dos objetivos e das questões de pesquisa; (3) elaboração da estratégia de busca e estabelecimento dos critérios de seleção; (4) determinação do método de pesquisa a ser aplicado; e (5) extração, sumarização e análise dos dados.

Na primeira etapa, o planejamento do protocolo da Revisão Sistemática da Literatura incluiu a definição dos procedimentos a serem seguidos para sua execução. Isso abrangeu a identificação da população e contexto, a elaboração dos critérios de inclusão e exclusão, e a determinação das fontes de pesquisa a serem utilizadas. Na segunda etapa, foram estabelecidos os objetivos específicos da pesquisa e formuladas as questões que orientariam a busca e a seleção preliminar dos estudos relevantes. Esta etapa foi crucial para assegurar que o foco da revisão fosse claramente delimitado e alinhado aos objetivos do estudo.

A terceira etapa envolveu a formulação da estratégia de busca e o estabelecimento dos critérios de seleção. Os termos-chave relevantes para o tema foram cuidadosamente definidos, juntamente com os parâmetros de inclusão dos estudos. Na quarta etapa, foi determinado o método de pesquisa a ser adotado, considerando o tipo de estudo a ser incluído, os critérios de qualidade a serem aplicados e a abordagem para a análise dos dados.

Por fim, na quinta etapa, os dados foram extraídos dos estudos selecionados, resumidos e analisados de forma sistemática. Essa análise teve como objetivo identificar padrões, tendências e lacunas na literatura existente sobre o tema.

4.2 Planejamento do Protocolo RSL

Nesta seção, é detalhada a etapa de planejamento e organização necessária para alcançar os objetivos da Revisão Sistemática da Literatura (RSL), abordando especificamente os procedimentos adotados para garantir a eficácia e a qualidade da revisão.

O protocolo da RSL desempenha um papel importante na criação de um processo estruturado para a condução das etapas da revisão, com o objetivo de otimizar o alcance dos objetivos estabelecidos. Para esta RSL, foi escolhido o *Parsifal* como ferramenta para

auxiliar na formalização, execução e síntese dos resultados. No *Parsifal*, foram definidos os objetivos, as questões de pesquisa, as palavras-chave, a *string* de busca, as fontes de pesquisa e os critérios de inclusão e exclusão dos estudos. O processo detalhado de planejamento estabeleceu uma base sólida para a condução de uma revisão abrangente e precisa da literatura relevante, conforme descrito nas seções 4.3 e 4.4.

4.3 Definição do Objetivo e Questões de Pesquisa

Nesta seção, é abordada a etapa de definição do objetivo da pesquisa, que constitui o propósito central deste estudo. Além disso, são delineadas as questões de pesquisa, que representam os elementos específicos a serem investigados e respondidos por meio de uma análise detalhada da literatura sobre o tema. Esta etapa é fundamental para orientar todo o processo de revisão, assegurando que as investigações sejam conduzidas de maneira direcionada e eficaz.

4.3.1 Objetivo

Nesta etapa, busca-se levantar dados e informações relacionados ao contexto da pesquisa e delinear as respectivas questões a serem investigadas. A formulação clara do objetivo é essencial para direcionar todas as ações subsequentes e assegurar que a pesquisa atinja seus propósitos de maneira eficaz.

Portanto, o objetivo desta Revisão Sistemática da Literatura (RSL) foi identificar tendências e avanços no uso de *Transfer Learning* na área da Educação. A intenção foi compreender como essa abordagem está evoluindo e sendo aplicada em diversos contextos educacionais. O foco esteve na análise de como o *transfer learning* possibilita o reaproveitamento do conhecimento documentado na literatura, contribuindo para o aprimoramento do processo de ensino-aprendizagem.

4.3.2 Questões de Pesquisa

As questões de pesquisa são essenciais para orientar a investigação e definir os parâmetros do estudo. Elas orientam a análise e oferecem uma estrutura para a coleta e interpretação dos dados, garantindo que a pesquisa seja conduzida de forma coerente e focada. Considerando o objetivo desta Revisão Sistemática da Literatura (RSL), foram

formuladas sete questões de pesquisa para estabelecer os critérios essenciais na seleção dos artigos e investigar as características identificadas nos problemas exploratórios. Assim, as seguintes *Questões de Pesquisa* (QP) foram definidas:

1. **QP1:** Qual é o estado atual da pesquisa sobre o uso de *transfer learning* na Educação?
2. **QP2:** Quais são os principais domínios educacionais (por exemplo, matemática, ciências, línguas) em que *transfer learning* tem sido aplicado com sucesso?
3. **QP3:** Quais são os conjuntos de dados mais frequentemente usados em pesquisas de *transfer learning* na Educação e quais são suas características?
4. **QP4:** Como *transfer learning* tem sido usado para melhorar o desempenho dos alunos em tarefas específicas, como classificação de textos, recomendação de conteúdo ou detecção de dificuldades de aprendizagem?
5. **QP5:** Quais são os desafios técnicos e práticos associados à implementação bem-sucedida de *transfer learning* em contextos educacionais?
6. **QP6:** Quais são as limitações e as preocupações éticas relacionadas ao uso de *transfer learning* na Educação, incluindo questões de privacidade e viés algorítmico?
7. **QP7:** Quais são as lacunas na pesquisa atual e as áreas que exigem mais investigação em relação ao uso de *transfer learning* na Educação?

4.4 Elaboração da *String* de Busca e definição de Critérios de Seleção

A elaboração da *string* de busca é um passo necessário para assegurar a abrangência e relevância dos estudos identificados durante a execução da Revisão Sistemática da Literatura (RSL). Este processo envolve a definição de palavras-chave e termos relacionados ao tema da pesquisa e ao contexto associado, que serão utilizados para recuperar artigos relevantes nas fontes de busca selecionadas.

Além disso, é necessário estabelecer critérios de seleção claros e objetivos para determinar quais estudos serão incluídos ou excluídos da revisão. Esses critérios garantem a qualidade e consistência dos estudos selecionados, contribuindo para a validade e confiabilidade dos resultados.

A seguir, são descritos os procedimentos utilizados para a formulação da *string* de busca e a definição dos critérios de seleção aplicados nesta pesquisa.

4.4.1 *String de Busca*

Para otimizar a busca por artigos relevantes e responder às questões de pesquisa, foram identificadas palavras-chave específicas para compor a *string* de busca. A *string* incluía termos relacionados ao contexto educacional, como "*Elementary School*,*Primary School*,*Secondary School*," e "*High School*," combinados com o conceito de *transfer learning*. Além disso, foram incorporados nomes de modelos de aprendizado de máquina amplamente utilizados, como *CNN*, *BERT*, *DenseNet*, *ResNet*, *GoogLeNet*, *VGGNet*, *MobileNet* e *Xception*. A construção da *string* foi orientada pela pergunta específica: "*Quais são os modelos de aprendizado de máquina mais utilizados na abordagem de transfer learning em diferentes níveis de ensino?*"

Além das categorias de termos mencionadas, foram incluídos sinônimos e variações linguísticas para cada termo, além de combinações entre eles. Por exemplo, termos como '*Elementary School*' também foram buscados como '*Primary School*' assim como diferentes formas de escrever nomes de modelos de aprendizado de máquina. Essa abordagem buscou garantir que a busca fosse abrangente, capturando uma maior diversidade de estudos relacionados ao tema.

Logo, a *string* de busca empregada para delimitar os trabalhos foi a seguinte:

("Elementary School"OR "Primary School"OR "Secondary School"OR "High School") AND "transfer learning" AND (CNN OR BERT OR DenseNet OR ResNet OR GoogLeNet OR VGGNet OR MobileNet OR Xception)

É relevante observar que a *string* de busca foi aplicada de forma consistente, sem variações, nas fontes de pesquisa definidas no protocolo desta RSL.

4.4.2 *CrITÉrios de Seleção*

Os critérios de seleção foram definidos para garantir a inclusão de estudos relevantes e a exclusão de trabalhos que não atendam aos objetivos da pesquisa. Foram considerados critérios de inclusão específicos: apresentar uma solução usando aprendizagem por transferência e apresentar uma solução considerando o contexto educacional.

Além disso, foram definidos critérios de exclusão para filtrar estudos que não atendem aos requisitos da pesquisa, tais como: artigos que não apresentam solução

usando aprendizagem por transferência, artigos que não apresentam solução relacionada ao contexto educacional, artigos relacionados ao público de estudantes do jardim de infância, artigos resumidos, capítulos de livros, documentos publicados na literatura cinzenta que não são considerados documentos formais, estudos secundários, e estudos terciários.

Estes critérios foram aplicados de forma sistemática e consistente durante o processo de seleção dos estudos, garantindo a qualidade e relevância dos artigos incluídos na revisão.

Para garantir a seleção adequada dos estudos, foram estabelecidos 02 (dois) critérios de inclusão e 08 (oito) critérios de exclusão, conforme apresentados na Tabela 1. Tais critérios serão aplicados durante as etapas de seleção preliminar dos estudos. Os critérios de seleção utilizados para delimitar os trabalhos foram os seguintes:

Tabela 1 – Critérios de Inclusão e Exclusão.

ID	Critérios de Inclusão
CI1	Apresentar uma solução usando aprendizagem por transferência.
CI2	Apresentar uma solução considerando o contexto educacional.
Critérios de Exclusão	
CE1	Estudos que não apresentam solução usando aprendizagem por transferência.
CE2	Estudos que não apresentam solução relacionada ao contexto educacional.
CE3	Estudos relacionados ao público de estudantes do jardim de infância.
CE4	Estudos resumidos.
CE5	Capítulos de livros.
CE6	Documentos publicados na literatura cinzenta, que não são considerados documentos formais.
CE7	Estudos secundários.
CE8	Estudos terciários.

Fonte: Autoria Própria (2024).

4.5 Método de Pesquisa

Nesta seção, são descritos os procedimentos e técnicas selecionados para atingir os objetivos propostos. Esta etapa é crucial para definir os métodos de coleta, análise e interpretação dos dados, assegurando a robustez e a validade dos resultados obtidos.

Os métodos de pesquisa são procedimentos sistemáticos fundamentais para a coleta e análise de dados, com o intuito de alcançar objetivos específicos. Para este estudo, a *string*

de busca foi executada nas seguintes bibliotecas digitais: *ACM Digital Library*¹, *IEEE Xplore*², *Scopus*³, *JSTOR*⁴, *PsycINFO*⁵, *PubMed*⁶, *Engineering Village (Compendex)*⁷, *ERIC (Education Resources Information Center)*⁸ e *Web of Science (WoS)*⁹.

4.6 Extração e Sumarização dos Dados, e Discussão dos Resultados

Nesta etapa, a partir dos estudos selecionados, após aplicação dos critérios de inclusão e exclusão, os dados acerca de cada estudo são extraídos e sumarizados para posterior análise. Este processo visa garantir que apenas os estudos de maior relevância e qualidade sejam considerados para análises mais aprofundadas, assegurando a robustez e a precisão dos resultados finais. Isso envolve a leitura completa dos estudos e organização das informações relevantes encontradas, incluindo dados demográficos, resultados de pesquisa, metodologias, técnicas e/ou ferramentas utilizadas e conclusões alcançadas.

A extração dos dados seguiu rigorosamente os campos definidos no protocolo da Revisão Sistemática da Literatura (RSL), assegurando consistência e precisão na coleta das informações. Tabelas e planilhas foram criadas para organizar os dados de maneira clara e acessível, facilitando a análise comparativa entre os estudos incluídos na revisão, além de possibilitar uma avaliação detalhada das variáveis e resultados obtidos.

Após a extração dos dados, conduz-se uma discussão detalhada dos resultados obtidos. Isso envolve a análise dos padrões e tendências identificados nos estudos revisados, além da avaliação criteriosa da qualidade e relevância das evidências encontradas. Essa análise permite não apenas identificar descobertas significativas, mas também contextualizá-las dentro do panorama atual da pesquisa, fornecendo uma base sólida para as conclusões e recomendações derivadas do estudo.

Resumidamente, a etapa de extração dos dados tem como objetivo coletar informações relevantes para a pesquisa, possibilitando uma análise abrangente do estado da arte em relação às questões de pesquisa investigadas. Neste estudo, após a execução da

¹ <http://dl.acm.org>

² <https://ieeexplore.ieee.org>

³ <http://www.scopus.com>

⁴ <https://www.jstor.org/>

⁵ <https://www.apa.org/pubs/databases/psycinfo/index>

⁶ <https://pubmed.ncbi.nlm.nih.gov/>

⁷ <https://www.engineeringvillage.com/>

⁸ <https://eric.ed.gov/>

⁹ <https://www.webofscience.com/>

string de busca nas fontes mencionadas e a aplicação dos critérios de inclusão e exclusão para a seleção dos trabalhos, identificou-se 31 estudos elegíveis para a etapa de extração de dados. Para esta etapa foi definido um conjunto de dados a serem extraídos, conforme apresentado na Tabela 2.

Tabela 2 – Listagem dos Dados para Extração dos Estudos Selecionados na RSL.

ID	Descrição do Dado
01	Título
02	Referência
03	Ano da publicação
04	Autores
05	País de origem dos autores do artigo ou para o qual o trabalho propõe solução
06	Principais objetivos do estudo
07	Metodologia usada no estudo (por exemplo, experimento, estudo de caso, análise quantitativa, análise qualitativa)
08	Principais métodos e técnicas empregadas
09	Contexto educacional em que o estudo foi conduzido (por exemplo, ensino fundamental, ensino superior, treinamento corporativo)
10	Conjuntos de dados utilizados
11	Origem dos conjuntos de dados
12	Tipo de Solução (por exemplo, <i>transfer learning</i> com CNN e SVM)
13	Como o <i>transfer learning</i> foi aplicado no estudo?
14	Técnicas utilizadas com a abordagem de <i>transfer learning</i>
15	Desempenho da solução proposta
16	Medidas usadas para avaliar o desempenho da solução proposta
17	Principais resultados ou descobertas do estudo
18	Forma de apresentação dos resultados (por exemplo, quantitativa ou qualitativa)
19	Como os resultados foram interpretados pelos autores?
20	Se a validação da solução ocorreu com potenciais usuários finais (ex: alunos, professores, educadores)
21	Breve descrição da solução
22	Limitações identificadas no estudo pelos autores
23	Como essas limitações podem afetar a interpretação dos resultados?
24	Principais conclusões do estudo
25	<i>Insights</i> ou recomendações específicas para educadores ou pesquisadores?

Fonte: Autoria Própria (2024).

Para assegurar a qualidade dos estudos selecionados e a pertinência de suas soluções para a pesquisa, foi delineado um processo de avaliação criterioso. Neste estudo, a qualidade foi medida pela capacidade dos trabalhos em abordar e/ou responder as questões de pesquisa estabelecidas para a RSL. Assim, na Tabela 3 foram elaboradas oito questões-chave para nortear essa avaliação de qualidade.

Essas oito questões foram essenciais para assegurar a qualidade e a consistência

Tabela 3 – Listagem de Questões para Avaliação da Qualidade dos Estudos na RSL.

ID	Questões de Qualidade (QQ)
QQ1	O estudo descreveu claramente o método de pesquisa utilizado para investigar o uso de <i>transfer learning</i> na Educação?
QQ2	A amostra utilizada no estudo foi representativa do contexto educacional em questão?
QQ3	A coleta de dados foi detalhadamente descrita?
QQ4	Os resultados do estudo foram apresentados de forma clara e concisa?
QQ5	O estudo reconheceu e discutiu as limitações metodológicas?
QQ6	O estudo discutiu a relevância prática dos resultados para a área da Educação?
QQ7	Os dados e os métodos foram descritos de maneira transparente?
QQ8	O estudo fez referência a outras pesquisas relevantes na área?

Fonte: Autoria Própria (2024).

dos estudos selecionados, proporcionando uma base sólida para a análise e interpretação dos resultados desta RSL.

4.7 Meta-análise

A meta-análise é uma ferramenta estatística poderosa e amplamente utilizada em pesquisas acadêmicas, particularmente em revisões sistemáticas da literatura. Ela permite a síntese e integração dos resultados de estudos individuais, oferecendo uma análise mais abrangente e robusta sobre um determinado tema, conforme discutido por Smith e Johnson (2019). Nesta seção, será examinado o processo de meta-análise como um componente essencial deste estudo, que investiga a aplicação de *transfer learning* na Educação.

Nesta pesquisa, a fase de meta-análise foi cuidadosamente estruturada para assegurar uma análise abrangente e precisa dos resultados dos estudos selecionados durante a execução da RSL. O processo foi dividido em etapas distintas, cada uma contribuindo para uma compreensão mais profunda do tema investigado. Inicialmente, foi realizado um estudo detalhado sobre os princípios e técnicas da meta-análise, estabelecendo uma base sólida para a execução dessa fase. Posteriormente, foram escolhidos *softwares* especializados que oferecem suporte eficaz para a meta-análise, garantindo uma abordagem metodológica consistente e confiável.

A definição dos modelos estatísticos foi um passo importante, envolvendo uma escolha criteriosa entre modelos de efeito fixo e aleatório, com base na natureza dos dados e nos objetivos da análise (GELMAN; HILL, 2006; RAUDENBUSH; BRYK, 2002). A heterogeneidade dos estudos foi rigorosamente avaliada, permitindo uma compreensão

mais precisa das variações nos resultados e a identificação de possíveis fontes de viés ou inconsistência, conforme discutido por Borenstein *et al.* (2011), Snijders e Bosker (2012). Finalmente, foram formuladas hipóteses norteadoras que orientaram a análise e a interpretação dos resultados, proporcionando descobertas valiosas sobre as relações entre as variáveis em estudo e os padrões observados nos dados.

Assim, a meta-análise foi realizada de maneira sistemática e rigorosa, com o objetivo de enriquecer o corpo de conhecimento existente sobre o uso de *transfer learning* na Educação. Este processo visa fornecer uma base sólida e relevante para a comunidade acadêmica e profissional.

4.7.1 Estudo sobre Meta-análise

Para aprofundar a compreensão do método de meta-análise e sua aplicação, foi realizado um estudo abrangente que abordou uma variedade de aspectos pertinentes. Este estudo incluiu a análise de pesquisas que exploraram o uso da meta-análise em diferentes contextos e disciplinas. Foram revisados guias práticos para a construção de uma meta-análise, além de estudos observacionais que aplicaram este método. Também foram investigadas questões relacionadas à interpretação de resultados, efeitos, heterogeneidade e testes de hipóteses associados à meta-análise.

A pesquisa bibliográfica foi realizada em bases de dados confiáveis, incluindo ACM, IEEE, Scopus e Google Acadêmico, visando identificar estudos que fornecessem evidências sobre a metodologia da meta-análise. O objetivo foi não apenas compreender os fundamentos teóricos por trás desse método, mas também explorar suas aplicações práticas e o impacto que poderia ter na pesquisa em questão.

Foi elaborado um mapa mental (ver Figura 9), que proporciona uma visão estruturada e organizada do método estatístico da meta-análise, aplicável a diversas áreas de pesquisa, como Medicina, Psicologia, Educação e Economia. Segundo Buzan e Buzan (2006), os mapas mentais são ferramentas poderosas para visualizar informações e fomentar a compreensão e a memória, facilitando a organização de ideias complexas de forma clara e concisa. No mapa, foram destacados os principais componentes da meta-análise, servindo como um guia estruturado, dividindo o processo de meta-análise em componentes-chave e passos essenciais. Cada elemento do mapa, desde a definição básica até a interpretação dos resultados, é cuidadosamente delineado para proporcionar uma

compreensão clara e concisa do método.

Este estudo possibilitou o desenvolvimento de estratégias e hipóteses específicas aplicadas na meta-análise apresentada neste trabalho. Essas estratégias englobaram a compilação e análise dos dados dos estudos selecionados, permitindo a realização de cálculos estatísticos que resultaram em uma medida sumária para cada estudo. Essa medida sumária forneceu a base para discutir as possíveis diferenças nos resultados dos estudos, contribuindo para uma compreensão mais aprofundada do tema em questão.

4.7.2 Identificação de Softwares

A identificação e seleção de *softwares* adequados para a condução da meta-análise foram cruciais para garantir a eficiência e a precisão do processo. Por meio de uma revisão da literatura especializada, foram identificados vários *softwares* amplamente utilizados para essa finalidade. De acordo com Higgins e Green (2019), o *RevMan*, desenvolvido pela *Cochrane Collaboration*, é destacado por sua integração com a metodologia *Cochrane* e sua aplicação em revisões sistemáticas na área da saúde. Segundo Borenstein *et al.* (2011), o *Comprehensive Meta-Analysis* (CMA) é amplamente adotado por sua capacidade de realizar análises estatísticas avançadas em meta-análises, oferecendo uma interface amigável e robusta para pesquisadores.

Além disso, a Linguagem e Ambiente *R* é frequentemente escolhida por sua flexibilidade e pela ampla disponibilidade de pacotes estatísticos. Entre esses pacotes, o *metafor* se destaca por sua aplicação específica em meta-análise, conforme descrito por Viechtbauer (2010). O *Stata*, amplamente reconhecido por sua capacidade de lidar com dados complexos, também é frequentemente utilizado em estudos de meta-análise. De acordo com StataCorp (2021), o *Stata* oferece ferramentas robustas e poderosas para análise estatística. Essas ferramentas são essenciais para analisar e sintetizar dados de estudos diversos, facilitando a interpretação dos resultados e a geração de descobertas significativas.

Os *softwares* citados oferecem uma variedade de recursos e funcionalidades que facilitam a organização, análise e interpretação dos dados obtidos dos estudos incluídos na meta-análise. No entanto, a escolha final do *software* utilizado foi realizada posteriormente, levando em consideração as necessidades específicas desta pesquisa e as características de cada uma das ferramentas identificadas.

Além de proporcionar suporte técnico para a execução da meta-análise, a seleção do *software* também teve um impacto significativo nos resultados e na abordagem adotada nesta pesquisa. Portanto, foi essencial realizar uma avaliação cuidadosa e criteriosa antes de determinar qual software seria utilizado, garantindo assim a qualidade e a confiabilidade dos resultados obtidos.

4.7.3 Definição do Modelo Estatístico

A escolha do modelo estatístico adequado é feita para garantir a precisão e a validade dos resultados obtidos na meta-análise. Diante da heterogeneidade dos estudos incluídos nesta pesquisa, que abordaram uma variedade de soluções, modelos e métricas distintas, foi determinado que o modelo de efeito aleatório seria mais apropriado para esta análise.

O modelo de efeito aleatório é recomendado quando há variação na verdadeira magnitude do efeito entre os estudos, o que é comum em pesquisas envolvendo diferentes contextos, métodos e populações. Este modelo não apenas considera a variabilidade dentro de cada estudo, mas também entre os estudos, possibilitando uma estimativa mais realista e abrangente do efeito de interesse, conforme descrito por Higgins e Green (2011).

Portanto, ao optar pelo modelo de efeito aleatório foi possível capturar e integrar a heterogeneidade dos estudos de forma mais adequada, fornecendo uma visão mais abrangente e representativa das tendências e padrões presentes na literatura revisada. Essa abordagem contribuiu para uma análise mais robusta e confiável dos resultados da meta-análise.

4.7.4 Avaliação da Heterogeneidade

A avaliação da heterogeneidade desempenha um papel fundamental na meta-análise, pois permite determinar a consistência dos resultados entre os estudos incluídos. Para realizar essa avaliação neste estudo, foram adotadas duas abordagens principais: o uso do gráfico de Floresta (do inglês, *Forest plot*) e a aplicação do teste de hipótese conhecido como Teste de Qui-Quadrado para Heterogeneidade (*Cochran Q-test*) de acordo com Higgins e Thompson (2002).

O *Forest plot* é uma ferramenta gráfica essencial em meta-análises, permitindo a visualização dos resultados de cada estudo individualmente, bem como dos intervalos

de confiança das medidas de efeito, como descrito por (BORENSTEIN *et al.*, 2009). A análise da dispersão dos resultados apresentados nos estudos fornece uma visão preliminar sobre a presença de heterogeneidade entre eles.

O Teste de Qui-Quadrado para Heterogeneidade (*Cochran Q-test*) é uma medida estatística amplamente utilizada em meta-análises para avaliar a variabilidade entre os estudos incluídos. De acordo com (EGGER *et al.*, 1997), esse teste compara a variabilidade observada nos resultados dos estudos com a variabilidade esperada apenas pelo acaso. O *Q-statistic* é calculado como a soma dos quadrados das diferenças ponderadas entre os resultados individuais dos estudos e a estimativa combinada de efeito. Um valor elevado da estatística Q sugere que a variabilidade entre os estudos é maior do que seria esperado por chance.

O *p-value* associado a esse teste nos permite avaliar a significância dessa variabilidade: um *p-value* baixo (geralmente < 0.05) indica que as diferenças entre os estudos são estatisticamente significativas, sugerindo a presença de heterogeneidade real. Isso significa que a variabilidade entre os estudos não é apenas devido ao acaso, mas reflete diferenças reais nos resultados. Por outro lado, um *p-value* maior sugere que a variabilidade observada pode ser explicada pelo acaso, indicando que as diferenças entre os estudos não são significativas e, portanto, os resultados podem ser considerados homogêneos. Esse teste é uma ferramenta essencial em meta-análises, pois fornece uma base estatística sólida para interpretar os resultados combinados de diferentes estudos, permitindo uma melhor compreensão da consistência dos achados.

Portanto, a combinação dessas abordagens permitiu uma avaliação abrangente da heterogeneidade nos resultados da meta-análise, fornecendo informações importantes sobre a consistência e a confiabilidade das conclusões obtidas.

4.7.5 Idealização de Testes de Hipóteses

A formulação dos testes de hipóteses tem o objetivo de investigar as características específicas identificadas nos estudos incluídos na meta-análise. De acordo com Hirakata *et al.* (2019), esses testes são uma ferramenta fundamental para constatar a autenticidade das afirmações obtidas a partir dos dados analisados. Para este trabalho, foram elaboradas afirmações relacionadas às pesquisas selecionadas, a fim de aceitar ou rejeitar hipóteses, com base em testes de hipóteses.

As hipóteses foram elaboradas considerando as tendências identificadas nos estudos incluídos na RSL. Essas hipóteses foram testadas com o objetivo de validar as conclusões obtidas a partir da meta-análise. Na Tabela 4 são apresentadas as hipóteses formuladas para este estudo.

Ao realizar os testes de hipóteses propostos, será possível verificar se as tendências identificadas nos estudos são estatisticamente significativas e se as conclusões obtidas são robustas o suficiente para embasar as recomendações a serem fornecidos ao final da pesquisa.

Tabela 4 – Idealização das Hipóteses.

ID	CENÁRIO 1
Hipótese Nula (H0)	A média das métricas de desempenho é a mesma para os três grupos.
Hipótese Alternativa (H1)	A média das métricas de desempenho é diferente entre os três grupos.
ID	CENÁRIO 2
Hipótese Nula (H0)	A efetividade dos modelos de aprendizado por transferência não varia significativamente entre os diferentes níveis de ensino.
Hipótese Alternativa (H1)	A efetividade dos modelos de aprendizado por transferência varia significativamente entre os diferentes níveis de ensino.

Fonte: Autoria Própria (2024).

4.8 Identificação de Lacunas

A identificação de lacunas na pesquisa é essencial para compreender os desafios não abordados ou áreas pouco exploradas em estudos anteriores. Conforme destacado por Lugoboni (2017), o primeiro passo para iniciar uma investigação é examinar estudos existentes e identificar questões não exploradas, gerando assim oportunidades para novas pesquisas.

De acordo com Queiroz *et al.* (2017), as lacunas de pesquisa abrem caminho para a continuidade de estudos anteriores, permitindo a avaliação das soluções propostas e a idealização de novas abordagens que possam preencher essas lacunas.

Neste estudo, propõe-se a elaboração de um fichamento utilizando planilhas para registrar as lacunas encontradas em cada estudo selecionado na RSL. Além disso, buscou-se propor ideias de soluções para os problemas identificados, oferecendo oportunidades de pesquisa que possam ser exploradas pela comunidade científica em trabalhos futuros. Essas

lacunas identificadas servirão como base para direcionar o desenvolvimento de estudos posteriores, contribuindo para o avanço do conhecimento na área investigada.

4.9 Considerações Finais do Capítulo

As etapas delineadas neste capítulo foram fundamentais para o sucesso da pesquisa, pois proporcionou uma estrutura clara e organizada para alcançar os objetivos propostos. A condução do estudo por meio de uma Revisão Sistemática da Literatura, seguindo passos bem definidos, desde o planejamento do protocolo até a extração dos dados e discussão dos resultados, permitiu uma abordagem transparente e abrangente.

Na etapa de meta-análise, a realização de um estudo sobre essa técnica, a definição de *softwares* de apoio e do modelo estatístico a ser empregado, bem como a avaliação da heterogeneidade dos estudos, foram passos essenciais para uma análise robusta e significativa dos dados.

A expectativa é que esse estudo gere resultados significativos, promovendo investigações e discussões sobre possíveis lacunas existentes na literatura. Essas lacunas identificadas podem servir como ponto de partida para novas pesquisas, contribuindo para o avanço do conhecimento na área e fornecendo *insights* valiosos para pesquisadores interessados em explorar temas relacionados.

5 RESULTADOS

Neste capítulo, são apresentados e discutidos os resultados da execução da Revisão Sistemática da Literatura e da Meta-análise como uma abordagem robusta e confiável para o levantamento do estado da arte da aplicação de Transferência de Aprendizagem na Educação.

5.1 Visão Geral dos Resultados da RSL

A execução da *string* de busca resultou em uma coleta significativa de estudos, como apresentado na Tabela 5, oferecendo uma visão geral do processo de seleção preliminar dos estudos. A busca foi realizada em nove bases de dados acadêmicas amplamente reconhecidas, abrangendo diferentes áreas do conhecimento, desde tecnologia da informação até ciências comportamentais e saúde mental.

Essa abordagem ampla e abrangente na busca por estudos relevantes para alcançar o objetivo deste trabalho permitiu a inclusão de uma variedade de perspectivas e contextos relacionados aos processos ou modelos de aprendizagem, centrados principalmente em estudantes e professores. A diversidade das bases de dados utilizadas também garantiu uma cobertura ampla da literatura relevante.

Os dados preliminares forneceram uma base sólida para a próxima etapa da pesquisa, que envolve uma análise mais detalhada dos estudos selecionados. Esta fase visa atingir os objetivos propostos e responder às questões de pesquisa delineadas no protocolo da Revisão Sistemática da Literatura (RSL).

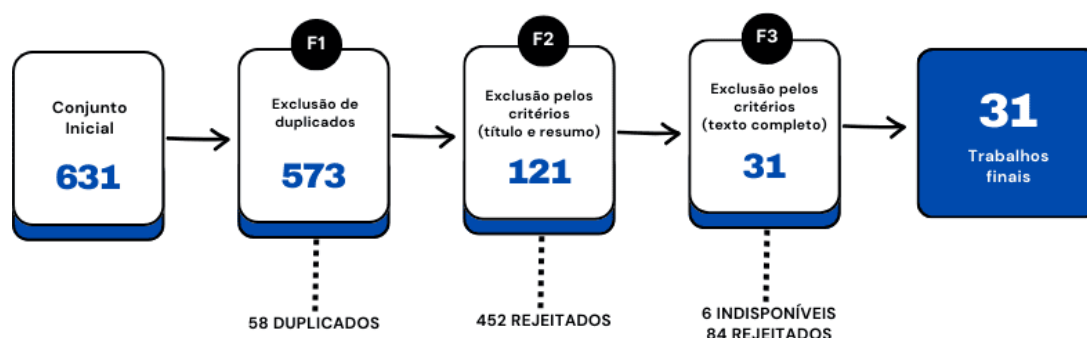
Tabela 5 – Síntese da Execução da *String* de Busca nas Fontes de Buscas.

Fonte	Número de estudos importados
ACM Digital Lybrary	163
ERIC	9
Engineering Village	36
IEEE Digital Library	18
JSTOR	8
PsycINFO	38
PubMed	3
Scopus	354
Web of Science	2
Total	631

Fonte: Autoria Própria (2024).

Os resultados da busca revelaram um total de 631 estudos selecionados após a execução da *string* de busca nas nove bases de dados acadêmicas. A distribuição desses estudos por fonte de busca foi a seguinte: 163 na *ACM Digital Library*, 9 na *ERIC*, 36 na *Engineering Village*, 18 na *IEEE Xplore*, 8 na *JSTOR*, 38 na *PsycINFO*, 3 no *PubMed*, 354 na *Scopus* e apenas 2 na *Web of Science*. Na Figura 3 é mostrada uma visão geral do processo de filtragem e seleção dos estudos.

Figura 3 – Visão geral do processo de seleção e extração de dados.



Fonte: Autoria Própria (2024).

Inicialmente, adotou-se uma abordagem de métodos mistos, conduzindo uma pesquisa em bases de dados orientada por uma *string* de busca Kitchenham e Charters (2007), resultando nos 631 estudos encontrados nas fontes supracitadas.

Posteriormente, foram realizados procedimentos adicionais utilizando a ferramenta *Parsifal*, incluindo a exclusão de estudos duplicados e a rejeição daqueles que não atendiam aos critérios de seleção definidos no protocolo da RSL. Foram identificados 58 estudos duplicados e 452 estudos foram rejeitados, resultando em 121 estudos selecionados para leitura completa e extração de dados.

Dentre esses 121 estudos, 84 estudos foram rejeitados por não aderir algum critério de inclusão e seis devido a indisponibilidade para *download*, impossibilitando a extração dos dados pelos pesquisadores, restando 31 estudos para a etapa de extração. Os estudos com acesso indisponível estão listados no Apêndice B e incluem: (i) Fung *et al.* (2020); (ii) Mohamed *et al.* (2021); (iii) Cooper e Sweller (1987); (iv) Saadna *et al.* (2022); (v) Sabuncuoğlu e Sezgin (2021); e (vi) Hamida *et al.* (2023).

Com o auxílio da ferramenta *Parsifal* para conduzir e gerenciar todas as etapas da RSL, desde a definição do protocolo até as etapas de seleção, extração de dados e avaliação

de qualidade, foi possível realizar uma análise minuciosa dos 31 artigos selecionados, como descrito no Apêndice A. Esse processo garantiu uma extração de dados detalhada para responder às questões de pesquisa propostas na revisão, além de permitir a identificação de lacunas e oportunidades de pesquisa.

A leitura cuidadosa de cada artigo proporcionou uma compreensão profunda dos conteúdos, facilitando a extração de informações relevantes para o estudo. Além disso, a utilização da ferramenta *Parsifal* contribuiu para uma gestão eficiente do processo, organizando os dados e as etapas de forma estruturada e sistemática.

Dessa maneira, o uso do *Parsifal* proporcionou uma abordagem metodológica robusta e transparente, garantindo a qualidade e a confiabilidade dos resultados obtidos ao longo da RSL. Conforme Figura 3, o processo de execução da RSL fornece uma compreensão clara das etapas envolvidas, facilitando uma compreensão do fluxo de trabalho adotado na pesquisa.

Na Figura 4 é apresentado uma representação visual das palavras-chave extraídas dos 31 artigos selecionados, utilizando a ferramenta *WordArt*¹. É uma maneira eficaz de destacar os termos mais frequentes e relevantes encontrados nos estudos analisados. É importante ressaltar que todos os estudos selecionados exploraram o conceito de *Transfer Learning* no contexto educacional, indicando a crescente importância desse tema de pesquisa integrando as áreas da Educação e Tecnologia.

Entre as palavras-chave mais recorrentes, pode-se observar termos como rede siamesa, processamento de linguagem natural, classificação de imagens, Inteligência Artificial na Educação, visão computacional, mineração de dados educacionais e análise de aprendizagem. Esses resultados destacaram áreas específicas de interesse e aplicação de *Transfer Learning* na Educação, refletindo a diversidade de abordagens e tecnologias utilizadas para promover a aprendizagem.

Além disso, é interessante notar a presença frequente de redes neurais convolucionais entre as palavras-chave extraídas dos 31 artigos. Isso sugere uma forte ênfase na utilização dessa arquitetura de rede neural como uma ferramenta eficaz para resolver problemas de aprendizado em diversos contextos educacionais. Essa observação revela as tendências e direções de pesquisa predominantes no campo de *Transfer Learning* aplicado à Educação, contribuindo para uma compreensão mais aprofundada das práticas e inovações

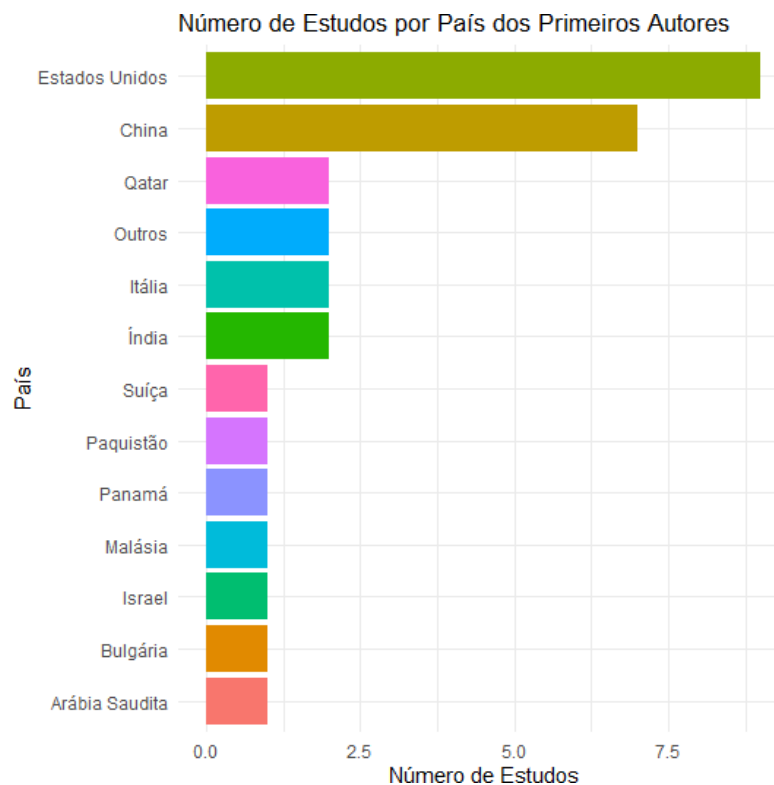
¹ <https://wordart.com/>

acervos digitais ou bibliotecas universitárias.

Neste estudo, realizou-se o mapeamento dos países de origem dos primeiros autores dos estudos selecionados. O objetivo dessa análise foi compreender os esforços de pesquisa relacionados à transferência de aprendizado no contexto educacional em diferentes partes do mundo. Ao examinar os dados, observou-se que a maioria dos artigos provém de periódicos e conferências específicas na área de Inteligência Artificial (IA) na Educação. Esse resultado não é surpreendente, dada a crescente importância da IA na transformação do setor educacional.

A utilização da IA oferece uma série de benefícios e oportunidades, incluindo a análise individualizada do desempenho e estilo de aprendizado de cada aluno, o que possibilita a personalização do conteúdo educacional de acordo com suas necessidades específicas. Isso, por sua vez, promove uma experiência de aprendizado mais adaptativa e eficaz, abrindo novas perspectivas tanto para os estudantes quanto para os professores. Na Figura 5 é apresentado o número de estudos por país, conforme a nacionalidade do primeiro autor de cada estudo.

Figura 5 – Distribuição dos Estudos por Nacionalidade do Primeiro Autor.



Fonte: Autoria Própria (2024).

Os Estados Unidos da América (EUA) emergem como uma potência significativa

em pesquisa em IA, com muitos dos principais centros de pesquisa e universidades do mundo localizados no país. Essa liderança se reflete em uma quantidade substancial de pesquisas e estudos, abrangendo diversas áreas, incluindo a aplicação do aprendizado por transferência na Educação. É notável que os EUA tenham liderado em termos de número de publicações, com 9 estudos identificados nesta pesquisa. Os autores desses estudos estão afiliados a instituições renomadas como o *Massachusetts Institute of Technology (MIT)*, *University of Oxford*, *Stanford University*, *Fremont High School* e *University of Colorado Boulder*. Muitos desses estudos exploraram uma variedade de técnicas de IA, como aprendizado de máquina, redes neurais, algoritmos genéticos, sistemas especialistas e processamento de linguagem natural. Os EUA é reconhecido por seus significativos investimentos em pesquisa, desenvolvimento e inovação em educação e tecnologia, o que contribui para sua posição proeminente no cenário mundial da IA na Educação.

A China emerge como o segundo maior país em termos de publicações, com um total de 7 estudos identificados nesta pesquisa. O país tem direcionado esforços consideráveis para melhorar seu sistema educacional e adotar tecnologias inovadoras como forma de aprimorar a qualidade do ensino. Os autores desses estudos estão afiliados a instituições como o *Guangzhou Morningsun Technology*, *Hubei Research Center for Educational Informationization*, *Huanggang High School of Hubei Province* e *School of Management*. Esses estudos empregaram uma ampla gama de métodos e técnicas para analisar a eficácia e a eficiência das pesquisas, refletindo o compromisso da China em promover a inovação educacional. O rápido crescimento econômico e tecnológico do país nas últimas décadas tem impulsionado um aumento significativo na pesquisa científica e tecnológica em diversas áreas, incluindo a Educação. Esse cenário reflete o compromisso da China em se tornar uma potência global em pesquisa e inovação.

Além dos Estados Unidos e da China, outros países também contribuíram significativamente para a pesquisa em transferência de aprendizado na Educação. Dois artigos foram identificados na Itália, com autores afiliados ao *Department of Mathematics, Computer Science and Physics* da *University of Udine* e ao *Politecnico di Milano*. No Qatar, também foram encontrados dois artigos, com autores vinculados ao *Department of Computer Science and Engineering* da *Qatar University* e ao *Qatar Computing Research Institute HBKU*, localizado em Doha, Qatar. Além disso, dois artigos foram atribuídos à Índia, com autores ligados ao *Computer Science and Engineering Department* da *Jawaharlal Nehru*

Technological University e ao *Department of CSE* da *Kongu Engineering College*. Essa distribuição geográfica diversificada reflete o alcance global da pesquisa em transferência de aprendizado na Educação e destaca a colaboração internacional nesse campo de estudo.

Além disso, houve contribuições individuais de outros países, como Israel, onde um estudo foi conduzido em uma escola em *Bat-Yam*, com o objetivo de desenvolver uma solução para a detecção precoce do Transtorno de Aprendizagem Específico em estudantes do ensino médio. Na Arábia Saudita, foi identificado um estudo realizado na *Faculty of Computing and Information Technology* da *King Abdulaziz University*, que, a partir da análise dos métodos e resultados apresentados, sugere-se que a solução proposta envolve o uso de técnicas de análise de dados e aprendizado de máquina para a detecção precoce da dislexia em crianças. No Paquistão, um estudo foi conduzido no *Department of Computer Science* da *National Textile University, Faisalabad*, e no *College of Computer and Information Sciences* da *Imam Mohammad Ibn Saud Islamic University (IMSIU)*, bem como no *Department of Computer Science* da *Quaid-i-Azam University, Islamabad*. Essa diversidade de localizações mostra como a pesquisa em transferência de aprendizado na Educação está sendo explorada em várias regiões do mundo, abordando questões educacionais específicas que são relevantes para cada contexto cultural e social.

Na Suíça foi identificado um estudo que consiste em uma revisão sistemática das abordagens recentes para a estimativa de dificuldade de perguntas em contextos educacionais, utilizando técnicas de Processamento de Linguagem Natural (PLN). No Panamá, um estudo utilizou um modelo pré-treinado/pré-processado em um grande volume de imagens. Na Malásia, o foco da pesquisa foi o desenvolvimento de uma solução de *transfer learning* para o reconhecimento de caligrafia para disléxicos, empregando uma Rede Neural Convolucional baseada na arquitetura conhecida por LeNet-5. Na Bulgária, foi realizado um estudo que avaliou o desempenho de vários modelos multilíngues pré-treinados, demonstrando o potencial desses modelos para melhorar a resposta a perguntas multilíngues. No entanto, em dois estudos, não foi possível identificar o país de origem dos autores. Essa ampla distribuição geográfica reflete o interesse global e a colaboração internacional na pesquisa sobre transferência de aprendizado na Educação, evidenciando a diversidade de abordagens e perspectivas em diferentes contextos ao redor do mundo.

Certamente, em alguns países, especialmente aqueles com recursos limitados em pesquisa e desenvolvimento, pode haver uma escassez de financiamento e recursos para

conduzir estudos avançados que envolvem técnicas complexas como *transfer learning*. Além disso, esses países podem ter diferentes prioridades de pesquisa, com alguns focando mais em questões específicas relacionadas à Educação e outros em áreas diferentes. Outro fator que pode influenciar é a disponibilidade limitada de conjuntos de dados educacionais acessíveis e adequados para pesquisa, o que pode ser considerado um sério impedimento. O acesso a dados de qualidade é fundamental para a aplicação bem-sucedida de técnicas de aprendizado de máquina. Esses desafios destacaram a importância de promover a colaboração internacional e o compartilhamento de recursos para garantir que a pesquisa em *transfer learning* na Educação seja inclusiva e abranja uma variedade de contextos e realidades ao redor do mundo.

A análise dos países dos primeiros autores dos estudos selecionados revelou uma distribuição geográfica diversificada, refletindo o interesse global em pesquisas relacionadas à transferência de aprendizado em diversos contextos educacionais. Os Estados Unidos emergiram como líderes nesse campo, demonstrando um forte compromisso com a inovação educacional e tecnológica. Com muitos dos principais centros de pesquisa e universidades situados no país, os Estados Unidos contribuíram significativamente para o avanço do conhecimento em IA na Educação.

Além disso, a presença da China como o segundo maior contribuinte é indicativa do rápido crescimento econômico e tecnológico do país, refletindo seu investimento crescente em pesquisa e desenvolvimento educacional. Outros países também apresentaram contribuições importantes para o corpo de conhecimento nessa área, cada um trazendo perspectivas únicas e abordagens inovadoras, como Itália Swamy *et al.* (2022) e Perseghin e Foresti (2023), Qatar Kunhoth *et al.* (2023) e Vachev *et al.* (2021), Índia Sharmila *et al.* (2023) e Bonthu *et al.* (2023), Israel Mor e Dardeck (2021), Arábia Saudita Alqahtani *et al.* (2023), Paquistão Ikram *et al.* (2023), Suíça Benedetto *et al.* (2023), Panamá Teran-Quezada *et al.* (2022), Malásia Rosli *et al.* (2021) e Bulgária Hardalov *et al.* (2020).

5.2.1.1 *Objetivos e Metodologias*

Ao investigar o desempenho da transferência de aprendizado, os pesquisadores podem identificar abordagens e estratégias mais eficazes na promoção da retenção e aplicação do conhecimento. A implementação de metodologias e modelos diferentes voltados para entender a transferência de aprendizado pode conduzir ao desenvolvimento

de abordagens inovadoras no campo da Educação. Isso inclui o uso de tecnologias avançadas, técnicas de aprendizado de máquina e abordagens pedagógicas que têm o potencial de melhorar significativamente a qualidade do ensino.

Compreender os desafios e obstáculos que podem ser enfrentados ao utilizar *transfer learning* na Educação é crucial para a criação de estratégias educacionais eficazes. Esses estudos podem identificar barreiras à transferência de aprendizado e orientar a criação de intervenções para superá-las. A análise dos dados dos estudos selecionados revelou uma variedade de soluções com diferentes focos, com proposta de implementação de metodologias e modelos para investigar o desempenho da transferência de aprendizado no contexto educacional.

Muitos estudos nesse campo envolveram o uso de modelos de aprendizado de máquina. Por exemplo, a pesquisa de Ding *et al.* (2019) apresentou uma abordagem que utilizou o aprendizado por transferência para melhorar o desempenho de previsão de modelos treinados a partir de dados históricos de cursos oferecidos anteriormente. Além disso, avaliou a eficácia da abordagem proposta em comparação com outras abordagens de aprendizado de transferência em MOOCs (do inglês, *Massive Open Online Courses*). A metodologia utilizada consistiu na aplicação de uma abordagem de aprendizado de transferência transdutiva automatizada baseada em representação de aprendizado com *auto-encoders*. A abordagem proposta compreendeu dois métodos de transferência alternativos: uma abordagem passiva que utilizou análise de componentes principais transdutiva e uma abordagem ativa que empregou uma função de perda de alinhamento de correlação. Esses métodos foram avaliados utilizando dados de cursos MOOCs e um problema de previsão de desistência.

A pesquisa de Huang e Guo (2021) teve como principais objetivos introduzir uma nova abordagem para avaliação de tópicos relevantes em escrita em inglês, utilizando aprendizado profundo e processamento de linguagem natural. Além disso, o estudo apresentou uma revisão da literatura sobre o uso de redes neurais convolucionais (do inglês, *Convolutional Neural Networks* - CNN) e redes neurais recorrentes (do inglês, *Recurrent Neural Network* - RNN) em tarefas de classificação de texto, e descreveu a metodologia e resultados dos experimentos realizados com o modelo *Siamese-Enhanced Representation through kNowledge IntEgration* (ERNIE). A metodologia utilizada no estudo baseou-se em experimentos nos quais os autores coletaram um conjunto de dados de cerca de 2500

ensaios de estudantes do ensino médio, rastreados aleatoriamente online. Esse conjunto de dados foi dividido em amostras positivas e negativas, e o modelo Siamese-ERNIE foi empregado para avaliar a relevância semântica entre o título e o conteúdo de cada ensaio. Além disso, os autores compararam o desempenho do modelo com outros modelos de aprendizado profundo, como CNN e RNN, em termos de precisão e velocidade de processamento, sendo os resultados dos experimentos apresentados e discutidos no estudo.

Mor e Dardeck (2021) realizaram um estudo para avaliar a capacidade do aprendizado profundo usando *transfer learning* e o modelo *MobileNet v2* para distinguir entre estudantes que têm um Transtorno Específico da Aprendizagem (do inglês, *Specific Learning Disability - SpLD*), tais como dislexia e disgrafia, e aqueles que não têm, com base em amostras de escrita à mão, e demonstrar a viabilidade de usar um algoritmo de aprendizado profundo para a triagem inicial de estudantes com SpLD usando *smartphones*. O estudo também destacou a importância da intervenção precoce em crianças com SpLD e sugeriu que um sistema como este pode contribuir significativamente para a detecção e intervenção precoces. A metodologia experimental adotada pelo estudo envolveu uma amostra de estudantes do ensino médio entre 15 e 18 anos de idade, nas séries 10 a 12, de uma escola em *Bat-Yam*, Israel. Utilizando o modelo *MobileNetV2* e *transfer learning*, o estudo treinou um algoritmo de aprendizado profundo para distinguir entre estudantes com e sem SpLD com base em amostras de escrita à mão. Os resultados foram avaliados usando medidas de desempenho, como precisão, sensibilidade e especificidade.

Benedetto *et al.* (2023) realizaram um estudo utilizando uma revisão sistemática das abordagens recentes para a estimativa de dificuldade de perguntas em contextos educacionais usando técnicas de Processamento de Linguagem Natural (PLN). O objetivo principal do estudo foi fornecer uma visão geral das abordagens existentes e avaliar sua eficácia em relação a diferentes tipos de perguntas e domínios educacionais. Além disso, o estudo também discutiu as limitações e desafios associados a essa tarefa e identificou possíveis direções futuras de pesquisa. A metodologia usada no estudo é uma revisão sistemática da literatura. Os autores realizaram uma pesquisa abrangente em várias bibliotecas digitais e mecanismos de busca para identificar estudos relevantes que abordassem a estimativa de dificuldade de perguntas em contextos educacionais usando técnicas PLN. Eles então analisaram criticamente os estudos selecionados e sintetizaram as informações relevantes em uma visão geral das abordagens existentes. O estudo não envolveu experimentos,

estudos de caso ou análises quantitativas ou qualitativas originais.

Com isso, nota-se que os estudos podem fornecer soluções de baixo custo, avaliar a eficácia da análise, ajudar crianças com dificuldades de aprendizagem específicas, desenvolver modelos de aprendizagem automática precisos e confiáveis que possam auxiliar na detecção e diagnóstico, discutir técnicas para melhorar o desempenho de modelos, desenvolver soluções baseadas em modelos de aprendizado de máquina, explorar o uso de *transfer learning* para prever o sucesso de algo relacionado à educação e, com isso, contribuir para melhorias na área da Educação. As metodologias apresentadas nesses estudos foram baseadas em experimentos, estudos de caso, análises quantitativas e qualitativas, assim como revisão sistemática da literatura para identificar estudos relevantes.

A análise dos estudos selecionados revelou uma variedade de abordagens inovadoras e promissoras no campo do *transfer learning* aplicado à Educação. Os resultados indicaram que essa técnica pode ser eficaz na melhoria do desempenho de modelos de aprendizado de máquina em uma variedade de contextos educacionais, desde a detecção precoce de transtornos de aprendizagem específicos até a estimativa de dificuldade de perguntas em avaliações educacionais.

Além disso, os estudos demonstraram a viabilidade e o potencial de soluções baseadas em modelos de aprendizado de máquina para auxiliar professores, alunos e profissionais da Educação em suas práticas e intervenções. Essas soluções não apenas ofereceram uma nova perspectiva para entender e abordar os desafios educacionais, mas também podem contribuir significativamente para a personalização do ensino, aprimorando a eficácia e a acessibilidade na Educação.

No entanto, é importante reconhecer que ainda existem desafios a serem superados, como a disponibilidade limitada de conjuntos de dados educacionais e a necessidade de garantir a interpretabilidade e a ética no desenvolvimento e uso de modelos de aprendizado de máquina na Educação. Portanto, são necessárias mais pesquisas e colaborações interdisciplinares para maximizar o potencial do *transfer learning* e promover avanços significativos na área da Educação.

5.2.1.2 Métodos e Técnicas

Na busca por aprimorar os métodos e técnicas empregados no campo do *transfer learning* aplicado à Educação, os pesquisadores têm explorado uma variedade de aborda-

gens inovadoras. Por meio da implementação de metodologias avançadas e técnicas de vanguarda, busca-se não apenas ampliar o entendimento sobre a transferência de aprendizado, mas também otimizar sua aplicação prática em contextos educacionais. Este texto explora algumas das principais metodologias e técnicas utilizadas em pesquisas recentes, destacando sua importância e potencial impacto na evolução do campo.

A pesquisa de Vachev *et al.* (2021) empregou métodos avançados que combinam informações ortográficas, lexicais, sintáticas e semânticas para gerar respostas em tarefas de questionamento e resposta. Adicionalmente, foram utilizados modelos baseados em entidades nomeadas e modelos neurais complexos para a geração de respostas candidatas. Essas abordagens visam aprimorar a precisão e a eficácia na geração de respostas, representando um avanço significativo na área de processamento de linguagem natural.

É interessante destacar que a pesquisa de Ding *et al.* (2019) explorou uma variedade de técnicas avançadas de aprendizado de máquina, incluindo aprendizado por transferência transdutiva automatizada e análise de componentes principais transdutiva. Essas abordagens são fundamentais para lidar com conjuntos de dados complexos e dinâmicos, como os dados de cursos MOOCs, e para prever com precisão a desistência dos alunos.

Essa variedade de métodos e técnicas utilizadas reflete a diversidade de abordagens que podem ser aplicadas ao *transfer learning* no contexto educacional. Ao adotar uma abordagem multidisciplinar e explorar diferentes estratégias, os pesquisadores podem melhorar continuamente a eficácia e a relevância das intervenções educacionais baseadas em *transfer learning*.

As técnicas de *transfer learning* desempenham um papel fundamental na otimização do uso de conjuntos de dados limitados, algo comum em ambientes educacionais. Ao aproveitar modelos pré-treinados em grandes conjuntos de dados, é possível transferir conhecimentos relevantes para tarefas específicas, mesmo diante de restrições de dados. Essa abordagem não apenas viabiliza a utilização eficiente de recursos, mas também permite a adaptação dos modelos pré-treinados a contextos específicos da Educação. Essa adaptação é crucial para assegurar que os conhecimentos transferidos sejam pertinentes e aplicáveis ao ambiente educacional, levando em consideração as características individuais dos alunos.

Um exemplo prático dessa aplicação é observado na pesquisa conduzida por Huang e Guo (2021), que empregou técnicas de aprendizado profundo e processamento de

linguagem natural. Nesse estudo, o modelo Siamese-ERNIE foi utilizado para avaliar a relevância semântica entre o título e o conteúdo de ensaios escritos por estudantes. Essa arquitetura de rede neural, baseada na técnica de máscara unificada, permite a aprendizagem de dependências de conhecimento e semânticas mais longas. Além disso, os autores realizaram uma comparação do desempenho do modelo Siamese-ERNIE com outros modelos de aprendizado profundo, como CNN e RNN, em termos de precisão e velocidade de processamento. A pesquisa também ofereceu uma revisão abrangente da literatura sobre o uso de redes neurais convolucionais e recorrentes em tarefas de classificação de texto, contribuindo assim para o avanço do conhecimento nesse campo específico.

O estudo conduzido por Mor e Dardeck (2021) empregou técnicas avançadas de aprendizado de máquina, destacando-se *Transfer Learning* e *Deep Learning*. O *Transfer Learning* é uma estratégia que consiste na reutilização de um modelo pré-treinado em uma tarefa específica como ponto de partida para um novo modelo a ser utilizado em outra tarefa. Já o *Deep Learning* é uma abordagem de aprendizado de máquina que se baseia no uso de redes neurais profundas para aprender a partir de dados. No contexto desse estudo foi utilizado o *MobileNetV2*, um modelo de rede neural convolucional pré-treinado, como ponto de partida para o modelo de aprendizado profundo.

Perseghin e Foresti (2023) empregaram uma variedade de métodos e técnicas para sua análise. Dentre elas, destaca-se o uso de CNN 2D, que foi empregada para processar imagens e vídeos, fornecendo uma abordagem robusta para a análise dos dados. Além disso, o estudo fez uso do *Transfer Learning*, que proporciona uma base sólida para a análise, aproveitando o conhecimento prévio do modelo e adaptando-o para o contexto específico do estudo. O conjunto de dados utilizado, denominado *Daily School Break*, foi especialmente criado para o estudo e contém vídeos originais gravados em um pátio de uma escola italiana. Essa escolha do conjunto de dados foi fundamental para garantir a relevância e a representatividade dos dados analisados. Para avaliar o desempenho dos modelos, a métrica de acurácia foi utilizada. Essa métrica permite medir a proporção de classificações corretas em relação ao total de classificações, fornecendo uma medida objetiva do desempenho dos modelos. Além disso, o estudo incluiu análises quantitativas e qualitativas dos resultados obtidos. A análise quantitativa envolveu uma análise estatística comparando a acurácia do modelo com diferentes arquiteturas de rede neural e conjuntos

de dados. Já a análise qualitativa foi descritiva, fornecendo *insights* sobre os resultados observados e suas implicações para o estudo. Essas abordagens combinadas permitiram uma avaliação abrangente e detalhada dos resultados obtidos pelos autores.

Os principais métodos e técnicas comumente encontrados nos estudos incluíram *Transfer Learning*, Processamento de Dados, Análise Quantitativa, Análise Qualitativa, Abordagens de Aprendizado de Máquina e Redes Neurais Convolucionais (CNNs). Essas técnicas desempenharam papéis fundamentais na pesquisa educacional, fornecendo ferramentas poderosas para análise e tomada de decisões informadas.

O Processamento de Dados é crucial para organizar e preparar informações educacionais para análise. Métodos eficazes de processamento de dados facilitam a extração de características relevantes e a aplicação de algoritmos de aprendizado de máquina, contribuindo para uma compreensão mais profunda dos dados. A Análise Quantitativa fornece medidas objetivas sobre o desempenho do *Transfer Learning* na Educação, permitindo uma avaliação precisa da eficácia das abordagens e a identificação de padrões estatísticos nos dados. Essa análise é essencial para fundamentar decisões baseadas em dados sólidos. Por sua vez, a Análise Qualitativa complementa a análise quantitativa, oferecendo uma compreensão mais profunda das experiências dos alunos, percepções dos educadores e outros fatores contextuais que podem influenciar o processo educacional.

As Abordagens de Aprendizado de Máquina são fundamentais para criar modelos preditivos, classificadores e sistemas adaptativos para a Educação. Essas técnicas permitem personalizar a educação, identificar padrões complexos nos dados e automatizar processos, promovendo uma aprendizagem mais eficiente e eficaz. Finalmente, as Redes Neurais Convolucionais (CNNs) são particularmente poderosas no processamento de dados visuais, como imagens. Em contextos educacionais, onde o conteúdo pode envolver elementos visuais, as CNNs são valiosas para extrair características complexas e representações eficientes, facilitando a análise de padrões visuais e reconhecimento de padrões em textos.

Essas técnicas e metodologias forneceram uma base sólida para a pesquisa educacional, permitindo a criação de modelos mais precisos, a compreensão aprofundada dos dados e a tomada de decisões fundamentadas. Ao combinar abordagens quantitativas e qualitativas, os pesquisadores podem obter uma visão holística dos desafios e oportunidades no campo da Educação. Com o contínuo avanço dessas técnicas e sua aplicação em ambientes educacionais, é possível promover uma aprendizagem mais adaptativa, eficiente

e inclusiva, capacitando alunos e educadores para alcançarem seu pleno potencial.

5.2.1.3 Contexto Educacional

Classificar os artigos sobre *transfer learning* na Educação de acordo com diferentes contextos educacionais oferece uma compreensão mais precisa de como essas pesquisas podem ser aplicadas em diversas etapas de ensino. Ao considerar etapas específicas, como o Ensino Fundamental I, II e Médio, bem como o Ensino Superior em cursos oferecidos de forma presencial ou pela Internet, os pesquisadores podem identificar abordagens e estratégias mais adequadas para cada ambiente educacional. Essa categorização permite uma análise mais aprofundada das necessidades, desafios e oportunidades em cada etapa de ensino, contribuindo para o desenvolvimento de intervenções mais eficazes e personalizadas para os alunos e educadores.

Essa análise visou oferecer uma compreensão mais precisa e específica da aplicabilidade do *transfer learning* em diferentes níveis educacionais. Ao considerar contextos variados, como o Ensino Fundamental, Médio e Superior, assim como cursos oferecidos pela Internet, foi possível identificar como essas técnicas podem ser adaptadas e implementadas de maneira mais eficaz em cada ambiente educacional. Cada etapa de ensino apresenta desafios e demandas específicas, e é importante entender como o *transfer learning* pode ser utilizado para enfrentar essas particularidades. Por exemplo, enquanto o estudo de Ding *et al.* (2019) focou em cursos online de Ensino Superior, como os MOOCs, a pesquisa de Mor e Dardeck (2021) explorou o potencial do *transfer learning* no Ensino Médio. Essa diversidade demonstra a versatilidade e relevância dessas técnicas em diferentes cenários educacionais.

Ao classificar os estudos de acordo com as diferentes etapas de ensino, foi possível direcionar a análise para a relevância dos métodos e resultados para públicos específicos, como crianças no Ensino Fundamental I, adolescentes no Ensino Fundamental II, jovens adultos no Ensino Médio e estudantes de Ensino Superior, considerando também as diferenças entre modalidades presenciais e online.

O estudo conduzido por Perseghin e Foresti (2023), embora realizado no contexto educacional de escolas públicas italianas, não especifica uma etapa de ensino específica. O objetivo principal do estudo foi desenvolver um sistema de detecção de violência em escolas públicas italianas, empregando uma CNN 2D e técnicas de *transfer learning*.

No contexto educacional, onde a segurança dos alunos é uma prioridade, a detecção de comportamentos violentos é de extrema importância. Este estudo destacou a necessidade de desenvolver conjuntos de dados específicos para a detecção de violência em ambientes escolares, propondo o conjunto de dados *Daily School Break*, composto por vídeos originais gravados em um pátio de uma escola italiana, como uma contribuição valiosa para essa área de pesquisa. Ao focar em contextos educacionais específicos, como escolas públicas italianas.

Entender como *transfer learning* pode ser aplicada em diferentes estágios da educação é fundamental para explorar maneiras específicas de melhorar a eficácia pedagógica, personalizando o ensino de acordo com as características e requisitos de cada etapa. O estudo conduzido por Ikram *et al.* (2023) ocorreu em um contexto de Ensino Superior, em um ambiente de sala de aula *offline*, onde os dados foram coletados durante palestras em duas turmas de uma universidade. Por outro lado, o estudo de Delaney e Bhatia (2021) foi realizado em salas de aula de matemática do oitavo ano, o que sugere que o contexto educacional foi o Ensino Fundamental II.

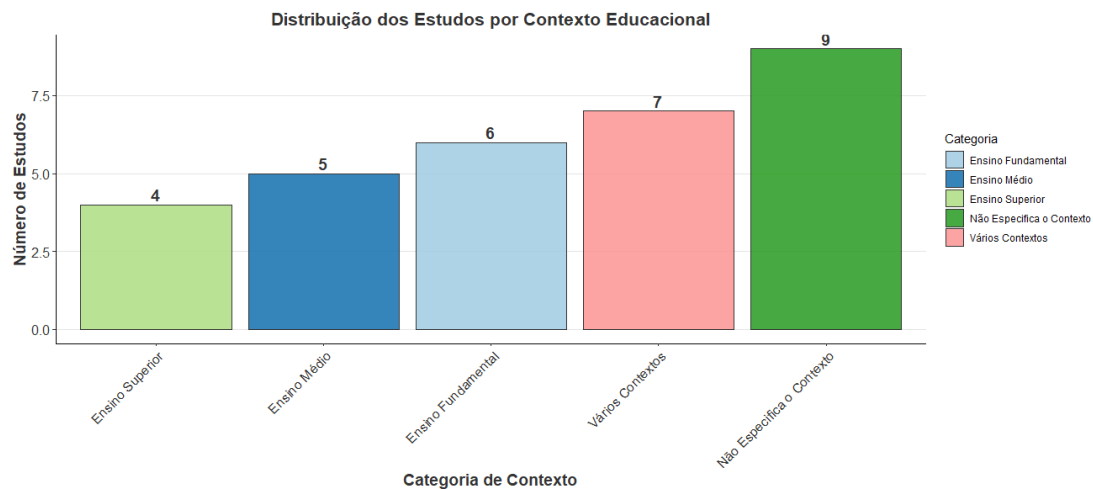
O estudo de Alqahtani *et al.* (2023) concentrou-se no desenvolvimento e avaliação de modelos de aprendizado de máquina para previsão e classificação de dislexia, utilizando diferentes tipos de dados, como amostras de caligrafia, exames de ressonância magnética e avaliações comportamentais e cognitivas. Os objetivos desse estudo foram melhorar a precisão e a confiabilidade do diagnóstico e previsão da dislexia, além de obter uma melhor compreensão dos mecanismos neurais e cognitivos subjacentes à dislexia. Portanto, esses estudos não se limitaram a um contexto educacional específico, mas visaram contribuir para um campo mais amplo de pesquisa e diagnóstico da dislexia, mostrando como o *transfer learning* pode ser aplicada de forma eficaz em diferentes contextos da Educação.

É interessante notar que alguns estudos não especificaram o contexto educacional em que foram conduzidos. Por exemplo, o estudo de Huang e Guo (2021) mencionou que o modelo foi treinado em um conjunto de dados com cerca de 2500 ensaios de estudantes do Ensino Médio, coletados aleatoriamente online, sem fornecer detalhes adicionais sobre o ambiente educacional específico. Da mesma forma, o estudo de Gillani *et al.* (2023) ofereceu uma visão geral das possibilidades e desafios da IA na Educação, além de uma estrutura para avaliar a ética da IA na Educação, sem especificar um contexto educacional particular. O estudo de Kunhoth *et al.* (2023), por sua vez, concentrou-se no

desenvolvimento de um sistema de diagnóstico preliminar baseado em imagens manuscritas para identificar a disgrafia em crianças, sem mencionar o contexto educacional em que o estudo foi conduzido. Essas lacunas na descrição do contexto educacional podem limitar a generalização dos resultados e a compreensão do impacto desses estudos em ambientes educacionais específicos.

A classificação dos estudos por etapas de ensino, conforme se observa na Figura 9, oferece informações valiosas para o desenvolvimento de ferramentas e recursos educacionais baseados em *transfer learning*, permitindo sua adaptação às necessidades específicas de cada faixa etária e contexto de aprendizagem. Essa abordagem possibilita a criação de aplicativos educacionais personalizados que atendam melhor às demandas e características de diferentes estágios educacionais, desde o Ensino Fundamental até o Ensino Superior. Além disso, ao identificar lacunas de pesquisa em termos de aplicação de *transfer learning* em cada etapa de ensino, essa categorização direciona futuras investigações para áreas específicas que ainda necessitam de exploração ou aprimoramento.

Figura 6 – Frequência dos 31 estudos da RSL por Etapa de Ensino.



Fonte: Autoria Própria (2024).

Dessa forma, a classificação por etapas de ensino não apenas amplia a compreensão sobre o impacto do *transfer learning* na Educação, mas também direciona o desenvolvimento de soluções mais eficazes e adaptáveis às diversas realidades educacionais.

Ao revisar os estudos sobre *transfer learning* em contextos educacionais, ficou evidente que essa abordagem está sendo aplicada em uma ampla variedade de cenários. Essa diversidade de aplicação reflete a natureza versátil e adaptável do *transfer learning*,

que pode ser ajustada para atender às necessidades específicas de diferentes públicos, disciplinas e métodos de ensino. A análise dos estudos por etapas de ensino revelou a importância de considerar as particularidades de cada fase do processo educacional ao implementar o *transfer learning*. Enquanto alguns estudos se concentram em estratégias para alunos do ensino fundamental, outros visam melhorar o ensino superior ou aprimorar a educação online. Essa segmentação por níveis educacionais permite uma abordagem mais direcionada e personalizada no desenvolvimento de soluções educacionais baseadas em *transfer learning*.

Além disso, ao categorizar os estudos por contexto educacional, também foi possível identificar lacunas de pesquisa e áreas que requerem mais investigação. Isso destaca a necessidade contínua de explorar e aprimorar o uso do *transfer learning* em diferentes contextos educacionais, garantindo que sua aplicação seja eficaz, ética e alinhada com as melhores práticas pedagógicas. Em última análise, os estudos revisados demonstraram que *transfer learning* tem o potencial de revolucionar a Educação, oferecendo oportunidades para melhorar a qualidade, acessibilidade e eficácia do ensino e aprendizagem em todo o mundo. Ao continuar a explorar e expandir seu uso, pode-se promover uma educação mais inclusiva, adaptável e centrada no aluno para as gerações futuras.

5.2.1.4 Conjuntos de Dados

A variedade no uso de diferentes conjuntos de dados e tamanhos na pesquisa sobre *transfer learning* na Educação pode ser atribuída a diversas razões. Nos estudos selecionados, observou-se uma ampla gama de conjuntos de dados, cada um adaptado para resolver problemas específicos. Essa diversidade reflete a complexidade do cenário educacional, que abrange uma variedade de disciplinas, tópicos e áreas de estudo.

Cada conjunto de dados foi cuidadosamente selecionado para corresponder às necessidades e características únicas de cada estudo. Por exemplo, alguns estudos se concentraram em problemas específicos, como detecção de dislexia ou previsão de desistência em cursos online, exigindo conjuntos de dados especializados para representar adequadamente esses contextos. Outros estudos exploraram a eficácia de abordagens de aprendizado em uma ampla gama de disciplinas, exigindo conjuntos de dados mais abrangentes e diversificados.

Além disso, a diferença nos tamanhos dos conjuntos de dados também pode ser

atribuída à disponibilidade de dados e à natureza específica dos problemas em questão. Em alguns casos, conjuntos de dados menores podem ser suficientes para demonstrar a viabilidade de uma abordagem de *transfer learning* em um contexto educacional específico. Em outros casos, conjuntos de dados maiores podem ser necessários para garantir resultados robustos e generalizáveis.

Em suma, a variedade de conjuntos de dados e tamanhos reflete a necessidade de adaptabilidade e personalização na pesquisa sobre *transfer learning* na Educação. Essa abordagem permite que os pesquisadores atendam às demandas específicas de cada problema educacional, garantindo que os modelos de *transfer learning* sejam capazes de generalizar efetivamente para diversas áreas do conhecimento e contextos educacionais.

A variedade nas etapas de ensino abordadas nos estudos sobre *transfer learning* na Educação reflete a importância de adaptar as abordagens de aprendizado às necessidades específicas de cada estágio educacional. Cada etapa de ensino apresenta características distintas, demandando conjuntos de dados que capturem adequadamente essas particularidades. Um exemplo elucidativo foi o estudo conduzido por Perseghin e Foresti (2023), que utilizou o conjunto de dados *Daily School Break*, desenvolvido especificamente para o problema tratado, sendo composto por vídeos originais gravados no pátio de uma escola italiana. É relevante ressaltar que a escassez de conjuntos de dados específicos para o reconhecimento e detecção de violência em ambientes educacionais motivou os autores a desenvolverem seu próprio conjunto de dados. Ao preencher essa lacuna, eles buscaram proporcionar uma representação mais precisa das ações violentas que ocorrem nas escolas públicas italianas. Além disso, o conjunto de dados foi projetado para ser leve e computacionalmente eficiente, o que o torna ideal para uso em aplicações sensíveis ao tempo ou em dispositivos de recursos limitados.

Por meio de técnicas de *transfer learning*, os autores refinaram o conjunto de dados, alcançando uma precisão de aproximadamente 95% no reconhecimento de ações violentas. Esse processo exemplifica como a adaptação de abordagens de aprendizado de máquina e o desenvolvimento de conjuntos de dados específicos podem contribuir significativamente para a eficácia das soluções educacionais baseadas em Inteligência Artificial. Essa abordagem sob medida reflete a compreensão da importância de considerar as características distintas de cada contexto educacional ao aplicar técnicas de *transfer learning*.

A disponibilização desse conjunto de dados para *download*, por meio de um endereço eletrônico fornecido no estudo, não apenas promove a transparência e a replicabilidade da pesquisa, mas também oferece uma oportunidade para que outros pesquisadores possam utilizar os mesmos dados para investigações futuras. Essa abordagem demonstrou o compromisso dos autores em fornecer recursos valiosos para a comunidade acadêmica e contribuir para o avanço do conhecimento no campo do reconhecimento e detecção de violência em ambientes educacionais.

Os conjuntos de dados utilizados no estudo de Swamy *et al.* (2022) sobre *Meta Transfer Learning for Early Success Prediction in MOOCs* consistem em interações de vídeo e quiz de alunos inscritos em 26 cursos de MOOCs ministrados por instrutores de uma universidade europeia entre 2013 e 2015. Esses conjuntos de dados foram totalmente anonimizados em relação às informações dos alunos e cobriram um total de 145.714 alunos. Após a remoção dos alunos que abandonaram precocemente os cursos, o conjunto de dados final continha 73.042 alunos. Esses conjuntos de dados foram essenciais para a análise do desempenho dos modelos de previsão de sucesso dos alunos com base em interações de vídeo e quiz em cursos online, permitindo uma avaliação realista da eficácia dos modelos propostos.

A pesquisa conduzida por Sharmila *et al.* (2023) empregou uma abordagem interessante na compilação do conjunto de dados final, utilizando dois conjuntos de dados separados para representar diferentes tipos de escrita manual: padrão e ilegível. Essa distinção foi fundamental para treinar e avaliar modelos de reconhecimento de caracteres manuscritos capazes de lidar com diferentes estilos de escrita. Para representar a escrita manual padrão, os pesquisadores extraíram dados do banco de dados *NIST Special Database 19*. Este banco de dados é uma fonte reconhecida e pública disponibilizada pelo *National Institute of Standards and Technology* (NIST) dos Estados Unidos. Ele contém uma vasta coleção de amostras manuscritas de escritores, incluindo imagens de caracteres isolados e formulários de amostra. Esses dados são acompanhados por classificações de verdade básica para as imagens e formulários, o que facilita a coleta e a análise de dados adicionais. Já para representar a escrita ilegível, foram utilizadas imagens de escrita manual espelhadas, proporcionando uma variedade de estilos de escrita desafiadores para o modelo reconhecer e interpretar. O conjunto de dados final compreendeu aproximadamente 2000 à 3000 imagens para cada letra, todas dimensionadas em 28x28 pixels. O

uso do *NIST Special Database 19* oferece uma base sólida e confiável para o treinamento e teste de modelos de reconhecimento de caracteres manuscritos, dada sua reputação e ampla utilização em pesquisas relacionadas. Essa escolha de conjunto de dados refletiu o compromisso dos pesquisadores em garantir a qualidade e a representatividade dos dados utilizados em sua pesquisa, contribuindo assim para avanços significativos no campo do processamento digital de imagens e reconhecimento de padrões.

É importante que os estudos mencionem claramente a fonte dos conjuntos de dados utilizados em suas pesquisas. Isso não apenas aumenta a transparência e a replicabilidade dos resultados, mas também permite uma avaliação mais precisa da validade e da qualidade dos dados empregados. No caso específico do estudo conduzido por Ding *et al.* (2019), a falta de menção à fonte dos dados dos *Massive Open Online Courses* (MOOCs) pode dificultar a compreensão completa do processo metodológico e a avaliação da representatividade dos dados utilizados. É crucial que os pesquisadores forneçam informações detalhadas sobre como os dados foram coletados, processados e preparados para análise, incluindo a descrição da base de dados original e quaisquer procedimentos de pré-processamento realizados.

A transparência na divulgação das fontes de dados também promove a confiança na pesquisa acadêmica e facilita a revisão por pares, permitindo que outros pesquisadores avaliem adequadamente a metodologia e os resultados apresentados. Portanto, é essencial que os autores sejam diligentes ao relatar a origem e as características dos conjuntos de dados utilizados em seus estudos, garantindo assim a integridade e a credibilidade da pesquisa científica.

A ausência de detalhes sobre a fonte específica do conjunto de dados utilizado no estudo de Huang e Guo (2021) pode representar uma limitação em relação à transparência e à replicabilidade do estudo. Sem informações claras sobre a origem e a qualidade dos dados utilizados, torna-se difícil para outros pesquisadores avaliarem a representatividade e a confiabilidade dos resultados obtidos. Idealmente, os autores deveriam fornecer informações detalhadas sobre a proveniência dos dados, incluindo a fonte original de onde foram coletados, os critérios de seleção dos dados e quaisquer procedimentos de pré-processamento aplicados. Isso não apenas aumentaria a transparência do estudo, mas também permitiria uma avaliação mais completa da validade dos resultados. Dessa forma, ao realizar futuras pesquisas, é importante que os pesquisadores sejam diligentes na

documentação e divulgação das fontes de dados utilizadas, garantindo assim a credibilidade e a integridade da pesquisa científica.

A falta de detalhes sobre conjuntos de dados específicos pode ser uma limitação significativa em estudos de *transfer learning* na Educação. Embora muitos estudos se concentrem na exploração de metodologias e algoritmos, a transparência e a replicabilidade da pesquisa podem ser comprometidas quando os detalhes sobre os conjuntos de dados utilizados não são fornecidos de forma adequada. No caso da pesquisa de Mor e Dardeck (2021), embora seja mencionado que amostras de escrita à mão foram coletadas dos estudantes, a falta de informações específicas sobre a origem dessas amostras, os critérios de seleção e quaisquer procedimentos de pré-processamento aplicados podem dificultar a avaliação da validade dos resultados e a comparação com outros estudos.

A heterogeneidade dos contextos educacionais pode representar um desafio significativo para a sistematização e padronização das informações sobre conjuntos de dados em estudos de *transfer learning* na Educação. Essa diversidade de cenários, desde diferentes etapas de ensino até variações nos métodos de coleta de dados, pode tornar difícil a comparação entre estudos e a generalização dos resultados.

A ausência de informações detalhadas sobre os conjuntos de dados em estudos de *transfer learning* na Educação ressalta a importância da qualidade e transparência dos relatórios nos estudos científicos. Ao fornecer detalhes completos sobre a origem, características e processamento dos conjuntos de dados utilizados, os pesquisadores podem melhorar a replicabilidade, a compreensão e a aplicabilidade de seus estudos. Isso também permite uma avaliação mais precisa da relevância e da eficácia das abordagens de *transfer learning* em contextos educacionais específicos.

Portanto, é fundamental que os pesquisadores adotem práticas de relatório transparentes e detalhadas, garantindo que as informações sobre os conjuntos de dados sejam documentadas de forma clara e precisa. Isso não apenas promove a credibilidade da pesquisa, mas também facilita a colaboração e a construção de conhecimento no campo do *transfer learning* aplicado à Educação.

5.2.1.5 Aplicação de Transfer Learning

O aumento do interesse e do desenvolvimento de pesquisas na área de Inteligência Artificial nos últimos anos é notável, refletindo o avanço constante no campo da IA e

suas aplicações em diversas áreas. O aprendizado por transferência tem se destacado como uma técnica fundamental nesse cenário, permitindo que modelos de IA aproveitem o conhecimento adquirido em uma tarefa e o apliquem a problemas relacionados, mas distintos.

Essa abordagem é especialmente poderosa porque permite que modelos pré-treinados, desenvolvidos com grandes conjuntos de dados e recursos computacionais significativos, sejam adaptados e aplicados a novos problemas com conjuntos de dados menores e recursos mais limitados. Isso acelera o processo de desenvolvimento de soluções em IA e aumenta sua acessibilidade para uma variedade de pesquisadores e profissionais.

Com o surgimento de novas pesquisas na área, é fundamental que esses estudos sejam disponibilizados em bibliotecas virtuais e repositórios online, facilitando o acesso e a disseminação do conhecimento para a comunidade científica. Essa prática não apenas promove a colaboração e o compartilhamento de ideias, mas também acelera o progresso na resolução de problemas complexos em diversas áreas do conhecimento.

O estudo de Ding *et al.* (2019) ilustrou o potencial de *transfer learning* para melhorar a eficácia do ensino e a qualidade da aprendizagem na Educação. Ao propor um método de aprendizagem por transferência baseado em representação, os autores demonstraram como o conhecimento adquirido em um curso pode ser aplicado de forma eficaz em diferentes contextos educacionais. Ao avaliar a transferibilidade dessa abordagem entre ofertas de cursos e entre cursos distintos, os pesquisadores destacaram a versatilidade e a adaptabilidade de *transfer learning* para lidar com uma variedade de situações e desafios educacionais. A utilização de uma série temporal por aluno, expressando as frequências de atividades MOOC por unidade de tempo, exemplifica a aplicação prática dessa técnica na análise de dados educacionais complexos. O estudo também destacou a importância da previsão de evasão como um exemplo de problema de modelagem preditiva relevante no contexto educacional.

O estudo de Huang e Guo (2021) destacou o uso eficaz do *transfer learning* na tarefa de classificação de similaridade de texto, utilizando o modelo *Siamese-ERNIE* como base. O *Siamese-ERNIE*, é uma variação do modelo *ERNIE (Enhanced Representation through kNowledge IntEgration)* desenvolvido pela Baidu, é uma rede neural pré-treinada capaz de capturar representações semânticas complexas em textos. Neste estudo, o *transfer learning* foi aplicado ao *Siamese-ERNIE*, inicialmente treinado em um conjunto de dados

específico, e posteriormente ajustado (*fine-tuned*) para a tarefa específica de classificação de similaridade de texto. O objetivo foi calcular a similaridade semântica entre o título e o conteúdo de um artigo, utilizando a distância de cosseno entre os *embeddings* semânticos dos dois textos. Os resultados do estudo demonstram que essa abordagem de *transfer learning* com o *Siamese-ERNIE* superou outras técnicas de aprendizado de máquina em termos de precisão e eficiência na classificação de similaridade de texto. Isso evidenciou a eficácia do *transfer learning* na criação de modelos de processamento de linguagem natural altamente precisos e adaptáveis, capazes de compreender e analisar a semântica de textos de forma sofisticada. Essa pesquisa destacou o potencial de *transfer learning* para impulsionar avanços significativos no campo da análise de texto e compreensão de linguagem natural.

O estudo realizado por Benedetto *et al.* (2023) destacou a diversidade de abordagens empregadas na estimativa da dificuldade de questões educacionais. Entre as técnicas exploradas estão Máquinas de Vetores de Suporte (SVMs), Modelos de Mistura Gaussiana (GMMs), Redes Neurais Convolucionais (CNNs) e Aprendizado por Transferência. Um exemplo interessante é o uso da Aprendizagem por Transferência para prever a dificuldade e o tempo de resposta de questões de múltipla escolha. Além disso, outra abordagem mencionada foi o emprego de SVMs com *embeddings word2vec* para estimar a dificuldade de questões baseadas em conhecimento de uma única palavra. O estudo destacou a importância da aplicação de uma variedade de técnicas e modelos no processo de estimativa de dificuldade de questões educacionais, demonstrando a eficácia e a relevância do aprendizado por transferência nesse contexto.

O estudo conduzido por Kunthoth *et al.* (2023) apresentou duas abordagens distintas para a classificação de imagens manuscritas, cada uma empregando técnicas diferentes de *transfer learning*. Na primeira abordagem, denominada *Transfer learning via fine-tuning*, foram desenvolvidas redes neurais convolucionais (CNNs) específicas para a tarefa de classificação de imagens manuscritas. Para isso, utilizaram a arquitetura DenseNet201 pré-treinada como ponto de partida. Essa arquitetura foi ajustada (*fine-tuned*) para se adaptar à tarefa específica de classificação das imagens manuscritas, permitindo que o modelo aproveitasse o conhecimento prévio adquirido durante o treinamento na tarefa original. Já na segunda abordagem, denominada Extração de características de CNN, as características das CNNs foram extraídas diretamente do conjunto de dados específico

da tarefa. Para isso, utilizaram uma rede DenseNet201 pré-treinada, que foi inicialmente treinada em um conjunto de dados grande e diversificado. Em seguida, as características aprendidas por essa rede foram extraídas das imagens do conjunto de dados da tarefa de classificação de imagens manuscritas. Essas características foram então utilizadas como entrada para o desenvolvimento de classificadores, empregando três algoritmos de aprendizado de máquina.

O estudo conduzido por Xiao *et al.* (2022) abordou uma variedade de modelos de aprendizado de máquina utilizados para a detecção de utilidade de *feedback*, destacando a aplicação do *transfer learning* para melhorar o desempenho dos modelos. Entre os modelos mencionados, estão modelos clássicos, como regressão logística, floresta aleatória e máquinas de vetores de suporte, bem como modelos de redes neurais, como text-CNN e LSTM. No entanto, uma abordagem específica se destacou: o uso de um modelo baseado em transformador, mais especificamente o *DistilBERT*, um modelo derivado do *BERT* (*Bidirectional Encoder Representations from Transformers*). O *DistilBERT* pré-treinado em grandes conjuntos de dados foi submetido a um processo de ajuste fino (*fine-tuning*) em um conjunto de dados menor, composto por comentários de *feedback*. Durante esse processo, o tokenizador também foi adaptado para se adequar a um novo vocabulário, proporcionando uma melhor representação dos dados específicos da tarefa.

O uso do *transfer learning* tem se mostrado uma estratégia promissora e eficaz em uma ampla gama de estudos na área de IA aplicada à Educação. Ao aproveitar o conhecimento adquirido em uma tarefa e aplicá-lo a outra tarefa relacionada, os pesquisadores têm conseguido melhorar significativamente a eficiência e a qualidade dos modelos de aprendizado de máquina. A análise dos estudos revisados revela uma diversidade de abordagens e técnicas de *transfer learning* aplicadas em diferentes contextos educacionais. Desde o uso de modelos pré-treinados em grandes conjuntos de dados até a adaptação desses modelos para tarefas específicas, como detecção de violência em escolas ou classificação de utilidade de *feedback*, o *transfer learning* tem demonstrado sua versatilidade e aplicabilidade em diversos cenários.

Além disso, os estudos destacaram a importância de considerar os conjuntos de dados específicos de cada contexto educacional, adaptando os modelos de acordo com as características e necessidades dos alunos e das instituições de ensino. Essa abordagem personalizada permite uma melhor generalização dos modelos e uma maior precisão nas

previsões e classificações. Em resumo, o uso de *transfer learning* nos estudos revisados tem sido fundamental para impulsionar a inovação e o avanço na área da Educação, fornecendo soluções mais eficientes e precisas para uma variedade de problemas e desafios educacionais. Essa abordagem continuará desempenhando um papel crucial na melhoria da qualidade do ensino e da aprendizagem, impulsionando ainda mais a pesquisa e o desenvolvimento de soluções inteligentes para a Educação.

5.2.1.6 Desempenho dos Modelos

A avaliação do desempenho dos modelos usando *transfer learning* em contextos educacionais diversos é essencial para garantir que esses modelos sejam eficazes e adequados para as tarefas específicas de aprendizagem. Para isso, uma variedade de métricas é comumente empregada, permitindo uma análise abrangente do desempenho dos modelos. Algumas dessas métricas incluem:

- **Acurácia (*Accuracy*):** mede a proporção de previsões corretas feitas pelo modelo em relação ao total de previsões. De acordo com Fawcett (2006), a acurácia é calculada como a soma de verdadeiros positivos e verdadeiros negativos dividida pelo total de previsões. Embora uma alta acurácia indique um bom desempenho do modelo, essa métrica pode ser enganosa em conjuntos de dados desbalanceados, onde outras métricas, como precisão e revocação, podem ser mais informativas;
- **Precisão (*Precision*) e Revocação (*Recall*):** são métricas usadas principalmente em problemas de classificação. De acordo com Fawcett (2006), a precisão mede a proporção de instâncias positivas previstas corretamente em relação ao total de instâncias previstas como positivas. Já a revocação mede a proporção de instâncias positivas previstas corretamente em relação ao total de instâncias positivas reais. Ambas as métricas são cruciais para avaliar a eficácia de um modelo de classificação, especialmente em conjuntos de dados desbalanceados;
- **F1-Score:** representa a média harmônica entre precisão e revocação, fornecendo uma medida balanceada entre as duas métricas. De acordo com Sokolova e Lapalme (2009), o F1-Score é especialmente útil quando se deseja uma avaliação balanceada do desempenho de um modelo de classificação, levando em consideração tanto a capacidade de identificar corretamente instâncias positivas quanto a de evitar falsos positivos;

- **Matriz de Confusão:** é uma tabela que mostra a frequência de classificações corretas e incorretas feitas pelo modelo para cada classe. De acordo com Powers (2011), a matriz de confusão é uma ferramenta essencial para avaliar o desempenho de modelos de classificação, fornecendo uma visão clara de como o modelo está realizando suas previsões em diferentes classes;
- **Área sob a Curva ROC (AUC-ROC):** é uma métrica que avalia a capacidade do modelo de distinguir entre classes positivas e negativas em um problema de classificação binária. De acordo com Fawcett (2006), a AUC-ROC é amplamente utilizada para medir o desempenho de classificadores, proporcionando uma avaliação robusta da capacidade discriminativa do modelo em diferentes limiares de decisão;
- **Erro Médio Quadrático (*Mean Squared Error* - MSE) ou Erro Médio Absoluto (*Mean Absolute Error* - MAE):** são métricas comumente usadas em problemas de regressão para avaliar a diferença entre os valores previstos e os valores reais. De acordo com Hastie *et al.* (2009), o MSE calcula a média dos quadrados dos erros entre os valores previstos e reais, enquanto o MAE calcula a média dos valores absolutos desses erros. Ambas as métricas são essenciais para quantificar a precisão de modelos de regressão em termos de erro de previsão;
- **Curva de Aprendizado:** é um gráfico que mostra como o desempenho do modelo varia com o tamanho do conjunto de dados de treinamento. Segundo Géron (2019), a curva de aprendizado é uma ferramenta essencial para entender como diferentes quantidades de dados de treinamento impactam na capacidade do modelo de generalizar para novos dados, ajudando a identificar se o modelo está sofrendo de viés alto (underfitting) ou variância alta (overfitting);
- **Métricas Específicas da Tarefa Educacional:** variam conforme a natureza da atividade educacional, abrangendo medidas adaptadas como taxa de aprovação, tempo de conclusão, taxa de retenção, entre outras. Segundo Siemens *et al.* (2013), a utilização de métricas personalizadas é crucial para uma avaliação precisa do sucesso e impacto de intervenções educacionais, proporcionando uma compreensão detalhada sobre o progresso dos alunos e a eficácia das estratégias pedagógicas;
- **Tempo de Treinamento e Inferência:** avalia o tempo necessário para treinar o modelo e para fazer previsões, sendo crucial para garantir a viabilidade prática do modelo em aplicações reais. Segundo Géron (2019), entender e otimizar o tempo

de treinamento e inferência são aspectos essenciais na construção de sistemas de aprendizado de máquina eficientes e escaláveis;

- **Feedback de Educadores e Estudantes:** além das métricas técnicas, o *feedback* qualitativo dos educadores e estudantes é crucial para avaliar a utilidade e a usabilidade dos modelos na prática educacional. Conforme Arnold e Pistilli (2012), integrar o feedback dos usuários finais é essencial para adaptar e melhorar continuamente as ferramentas de aprendizado automático, garantindo que atendam efetivamente às necessidades educacionais específicas.

Essas métricas em conjunto proporcionam uma visão abrangente do desempenho dos modelos e sua adequação para as necessidades educacionais, permitindo ajustes e melhorias contínuas.

Para demonstrar o desempenho das soluções utilizando as métricas mencionadas, foram empregados diversos métodos de aprendizado por transferência, CNNs, o modelo Siamese-ERNIE e outras arquiteturas de IA. Entre essas arquiteturas estão incluídas RNNs, *transformers*, algoritmos de aprendizado profundo, máquinas de vetores de suporte (SVMs), redes neurais artificiais (RNAs), eliminação de características recursivas com validação cruzada (RFE-CV), deep learning (DL) e árvores de decisão (DTs, do inglês *Decision Tree*). A avaliação do desempenho desses métodos e modelos foi realizada utilizando métricas específicas, como acurácia, precisão, revocação, *F1-Score*, área sob a curva ROC (AUC-ROC), erro médio quadrático (MSE), erro médio absoluto (MAE), entre outras, para fornecer uma compreensão abrangente do quão eficazes são essas soluções na resolução de problemas educacionais.

No estudo de Mor e Dardeck (2021), para avaliar o desempenho do modelo de aprendizagem profunda, os pesquisadores utilizaram cinco métricas principais: área sob a curva (AUC), precisão, *recall*, *F1-Score* e acurácia. Essas métricas foram calculadas com base nos resultados obtidos no conjunto de validação, composto por 50 imagens de caligrafia. De acordo com os resultados apresentados, o modelo alcançou os seguintes valores de métricas: AUC = 0,89, precisão = 0,94, *recall* = 0,89, pontuação *F1-score* = 0,91 e acurácia = 0,92. Esses resultados sugerem que o modelo demonstrou um alto nível de precisão na detecção de Distúrbios de Aprendizagem Específicos (SpLD) com base em amostras de caligrafia.

Na pesquisa de Perseghin e Foresti (2023), para avaliar o desempenho do modelo

de aprendizagem por transferência, os autores utilizaram a precisão como métrica principal. Além disso, eles empregaram precisão, *recall* e *F1-Score* para comparar os resultados obtidos com arquiteturas de baixo custo. Essas métricas são amplamente reconhecidas e utilizadas em tarefas de classificação para avaliar o desempenho do modelo. O estudo forneceu uma análise detalhada dos resultados obtidos e os comparou com outros trabalhos relacionados na área. Os resultados experimentais demonstraram que, ao refinar o conjunto de dados por meio de *web scraping* e aplicar técnicas de transferência de aprendizagem, a classificação da rede alcançou uma precisão em torno de 95%. No entanto, o estudo também mencionou que alguns desafios persistem, como o ruído nas imagens e a dificuldade de obtenção de imagens com crianças como sujeitos principais. No geral, o sistema apresentou resultados promissores e pode ser aprimorado ainda mais, seja analisando a série temporal de quadros ou implementando áudios originais usando o conjunto de dados *Daily School Break dataset* (DSB), que consiste em vídeos originais gravados em um pátio de uma escola italiana durante o intervalo escola.

Na pesquisa de Ding *et al.* (2019) foi utilizado a métrica de área sob a curva característica de operação do receptor (AUC) para medir o desempenho de todos os modelos preditivos. As pontuações AUC foram selecionadas a partir da arquitetura de IA com melhor desempenho e foi calculada a média de todos os pares possíveis de cursos de origem e de destino em todas as semanas. Os autores realizaram comparações e contrastes entre os cursos usados para avaliar os métodos de aprendizagem por transferência e demonstraram que os métodos propostos baseados na aprendizagem de representação melhoraram a previsão para o curso alvo e eliminaram a necessidade de engenharia manual de recursos. As pontuações AUC de previsão de abandono melhoraram em 8% usando qualquer um dos métodos em comparação com a linha de base da transferência ingênua. Os autores descobriram que seus modelos transferidos superaram consistentemente outras linhas de base de transferência e alcançaram desempenho semelhante em comparação com os modelos de rótulo-verdade (sem transferência) que foram treinados no curso alvo. Nesse sentido, os autores acreditam ter feito progressos significativos na resolução do problema de transferência de aprendizagem em MOOCs.

Na pesquisa de Xiao *et al.* (2022), o desempenho do modelo de *transfer learning* foi avaliado utilizando a métrica *F1-Score*, que combina precisão e *recall*. O estudo comparou o desempenho do modelo resultante com outros modelos, incluindo modelos

de aprendizado de máquina clássicos e modelos de rede neural, em várias tarefas de classificação de características de *feedback*. O desempenho do modelo foi avaliado em relação a cada uma dessas tarefas, e o F1-Score foi calculado para cada modelo em cada tarefa. O modelo resultante, que utilizou *transfer learning* com o modelo BERT, obteve um *F1-Score* de 0,68 na tarefa de detecção de utilidade de *feedback*. Embora esse desempenho tenha sido inferior ao de outros modelos em algumas tarefas, o estudo mostrou que o uso de *transfer learning* pode ser uma técnica eficaz para melhorar o desempenho de modelos de aprendizado de máquina em tarefas específicas.

Adaptar as métricas de desempenho ao contexto específico da tarefa educacional é fundamental para garantir uma avaliação precisa e relevante dos modelos. As métricas escolhidas devem estar alinhadas com os objetivos do estudo e as características dos dados, levando em consideração as implicações práticas da aplicação do modelo na prática educacional.

Além disso, a interpretação dessas métricas é crucial para entender o impacto e as necessidades específicas do contexto educacional. Por exemplo, métricas como precisão, *recall* e *F1-Score* são importantes para avaliar a capacidade de um modelo em identificar corretamente instâncias positivas e negativas em problemas de classificação. No entanto, em determinados contextos educacionais, outras métricas podem ser mais relevantes, como taxa de aprovação, tempo de conclusão, taxa de retenção, entre outras.

Portanto, ao realizar estudos que envolvam a aplicação de modelos de aprendizado de máquina na Educação, é essencial escolher e interpretar as métricas de desempenho de forma cuidadosa e criteriosa, garantindo que os resultados obtidos sejam significativos e relevantes para o contexto educacional em questão.

5.2.1.7 Principais Resultados e Descobertas

A interpretação dos resultados e descobertas ao utilizar *transfer learning* em contextos educacionais é crucial para compreender o impacto e a eficácia dessa abordagem. Ao analisar os resultados, é importante considerar os objetivos específicos da pesquisa e a natureza das tarefas educacionais em questão. Uma prática comum é comparar os resultados obtidos com modelos de *transfer learning* com *baselines*, que podem incluir modelos treinados do zero ou modelos que não incorporam *transfer learning*. Essa comparação permite avaliar a eficácia do uso de *transfer learning* em relação a abordagens tradicionais.

Além disso, ao interpretar os resultados, é fundamental considerar não apenas as métricas de desempenho, mas também o contexto educacional em que o modelo é aplicado. Por exemplo, um modelo pode ter um desempenho superior em termos de métricas tradicionais, como acurácia ou *F1-Score*, mas pode não ser adequado para o contexto educacional devido a outros fatores, como interpretabilidade, adaptabilidade ou viabilidade prática.

Analisar os resultados em busca de padrões é uma abordagem fundamental para compreender a eficácia do *transfer learning* em contextos educacionais. Ao examinar os resultados dos estudos que utilizam essa técnica, é possível identificar tendências valiosas que podem informar futuras pesquisas e práticas educacionais. A observação de que 24 estudos apresentam análises quantitativas utilizando métricas como acurácia, precisão, *recall* e *F1-Score* é significativa, pois fornece uma base consistente para comparar o desempenho dos sistemas em diferentes contextos educacionais. Essas métricas permitem uma avaliação objetiva do desempenho dos modelos de *transfer learning* e facilitam a comparação entre diferentes abordagens e estudos.

Além disso, o fato de que a maioria dos estudos forneceu uma análise detalhada dos resultados e comparações com outros trabalhos relacionados na área é muito promissor. Isso sugere um alto nível de rigor científico e uma abordagem sistemática na avaliação do desempenho do *transfer learning* na Educação. Ao identificar características relevantes para diferentes etapas de ensino, estilos de aprendizado ou outras variáveis, os pesquisadores podem oferecer informações valiosas para o desenvolvimento de sistemas educacionais mais eficazes e personalizados. Esses *insights* podem informar a criação de abordagens de ensino mais adaptativas, o desenvolvimento de recursos educacionais personalizados e a implementação de estratégias de apoio ao aprendizado mais eficientes.

É interessante observar a variedade de abordagens adotadas pelos estudos na apresentação de seus resultados. Enquanto alguns optam por uma análise quantitativa detalhada, utilizando métricas específicas para avaliar o desempenho dos modelos propostos, outros complementam essa análise com uma discussão qualitativa mais ampla. Dos estudos mencionados, dois optam por uma abordagem predominantemente qualitativa, enfatizando discussões e comparações entre diferentes abordagens e resultados relatados na literatura. Essa abordagem pode oferecer uma visão mais holística do impacto e das implicações dos modelos propostos.

Por outro lado, quatro estudos combinaram tanto análises quantitativas quanto qualitativas em sua apresentação de resultados. Ao utilizar medidas de correção como exatidão, precisão, *recall* e *f1*, esses estudos forneceram uma avaliação abrangente do desempenho dos modelos propostos. Além disso, ao ilustrar características dos alunos e resultados dos experimentos por meio de figuras e tabelas, eles ofereceram uma compreensão mais visual e detalhada dos resultados. No entanto, é importante notar que um estudo não apresentou dados quantitativos ou qualitativos, optando por fornecer um quadro para os educadores avaliarem a aplicabilidade e os riscos dos sistemas de IA em seus próprios contextos. Embora essa abordagem possa oferecer uma perspectiva importante sobre considerações éticas e práticas, a falta de dados empíricos pode limitar a capacidade de validar as conclusões e recomendações do estudo.

Em suma, a diversidade de abordagens na apresentação de resultados reflete a complexidade e multidimensionalidade das questões relacionadas ao uso de *transfer learning* na Educação. Cada abordagem tem seus próprios pontos fortes e limitações, e uma combinação de análises quantitativas e qualitativas pode oferecer uma compreensão mais completa e informada do impacto dessas tecnologias em contextos educacionais.

A interpretação dos resultados realizada pelos autores de Ding *et al.* (2019) destacou a eficácia dos métodos de aprendizado de transferência online baseados em aprendizado de representação na melhoria da previsão de desistência em MOOCs. Eles observaram que esses métodos podem ser especialmente benéficos para melhorar o desempenho de previsão em grupos minoritários, além de facilitar o fornecimento de suporte de aprendizado mais específico. Um dos pontos-chave levantados pelos autores foi a constatação de que eventos relacionados a vídeo mostram ser mais transferíveis e preditivos do que outros eventos. Essa descoberta ressaltou a importância de considerar a natureza dos dados em *transfer learning*, destacando que certos tipos de eventos podem ter uma influência maior no desempenho dos modelos. No contexto geral, os autores concluíram que a abordagem proposta tem o potencial de melhorar a eficácia dos sistemas de suporte de aprendizado em MOOCs e outros ambientes de aprendizado online. Isso sugere que a aplicação de técnicas de *transfer learning* pode desempenhar um papel significativo na personalização e otimização da experiência de aprendizado dos alunos em ambientes educacionais digitais.

A pesquisa conduzida por Perseghin e Foresti (2023) evidenciou a eficácia do sistema proposto na detecção de ações violentas no conjunto de dados *Daily School Break*,

alcançando uma precisão de aproximadamente 95%. Os autores destacaram que o uso de técnicas de transferência de aprendizagem, juntamente com o refinamento do conjunto de dados por meio de *web scraping*, contribuiu significativamente para a melhoria da precisão do sistema. A análise detalhada dos resultados e a comparação com outros trabalhos relacionados na área fortalecem a confiança na solução apresentada. No geral, o estudo ofereceu uma solução viável e de baixo custo para a detecção de ações violentas em escolas públicas, destacando sua relevância para a segurança e o bem-estar dos alunos.

Da mesma forma, o estudo conduzido por Huang e Guo (2021) revelou que o modelo Siamese-ERNIE obteve um desempenho satisfatório na tarefa de avaliar a similaridade semântica entre títulos e conteúdos de artigos, com uma precisão de 62% e um NDCG@4 de 0,72. A comparação do desempenho do modelo com diferentes tamanhos de camada MLP demonstrou que a precisão e o *Normalized Discounted Cumulative Gain* (NDCG) uma métrica crucial em sistemas de recomendação e recuperação de informações para avaliar a qualidade de uma lista classificada de documentos, diminuem à medida que o tamanho da camada MLP diminui. Esses resultados indicam que o modelo Siamese-ERNIE apresenta uma abordagem promissora para avaliar a relevância de tópicos em tarefas de escrita em inglês, sugerindo sua aplicabilidade e utilidade em diferentes contextos educacionais.

5.2.1.8 Conclusões e Validações

Nota-se, após a revisão dos 31 estudos selecionados nesta RSL, que os procedimentos de validação e as métricas de desempenho variam consideravelmente entre os estudos, refletindo a diversidade de metodologias, modelos e cenários de aplicação. Enquanto alguns estudos forneceram detalhes abrangentes sobre os métodos de validação e métricas utilizadas, outros não expuseram essas informações de forma explícita. No entanto, mesmo diante dessa diversidade, é possível observar que os estudos que forneceram uma análise abrangente do desempenho de seus modelos tendem a demonstrar melhorias significativas ao utilizar os métodos propostos. Essas melhorias indicaram a eficácia dos modelos implementados e reforçaram a validade das abordagens de *transfer learning* em contextos educacionais.

O estudo de Gillani *et al.* (2023) destacou uma perspectiva fundamental sobre a colaboração entre pesquisadores e profissionais da educação no contexto da IA. Embora o

estudo não forneça detalhes sobre a validação da solução com potenciais usuários finais, como estudantes, professores ou educadores, ele enfatizou a importância dessa colaboração para uma compreensão mais profunda e holística da aplicação da IA na Educação. A ênfase na colaboração entre investigadores e profissionais da educação reflete o reconhecimento da complexidade da implementação de soluções de IA na área. Essa colaboração não apenas permite que os educadores compreendam melhor os aspectos técnicos da IA, mas também capacita os pesquisadores a desenvolverem soluções mais alinhadas com as necessidades e realidades do campo da Educação.

O estudo de Ding *et al.* (2019) destacou a importância da aprendizagem por representação e por transferência na melhoria da previsão de abandono em MOOCs e em outros ambientes de aprendizagem online. Ao utilizar dados de fluxo de cliques de vários cursos, os autores desenvolvem uma abordagem que aprende uma representação de recursos preditivos comuns a diferentes MOOCs. Os resultados indicaram que essa abordagem supera outras técnicas de aprendizagem por transferência e uma linha de base de transferência ingênua. Embora não haja informações específicas sobre a validação da solução com potenciais usuários finais, como alunos, professores ou educadores, os autores sugerem que a abordagem proposta pode melhorar a eficácia dos sistemas de apoio à aprendizagem em MOOCs e em outros ambientes de aprendizagem online. Essa sugestão implica que a solução pode ser relevante e útil para os usuários finais, destacando a importância da aplicabilidade prática da pesquisa em contextos educacionais reais.

O estudo de Bian *et al.* (2019) abordou a inferência de emoções acadêmicas a partir de expressões faciais espontâneas em um ambiente de aprendizagem online. Eles demonstraram que a base de dados de expressões faciais espontâneas pode ser fundamental para o desenvolvimento de algoritmos de inferência de emoções nesse contexto. Os resultados revelaram que a precisão da inferência de emoções acadêmicas varia de acordo com a emoção em questão e que combinar várias emoções pode aumentar a precisão da inferência. Para realizar o estudo, os participantes foram convidados a assistir a vídeos de cursos online enquanto suas expressões faciais eram gravadas. O objetivo era inferir emoções acadêmicas a partir das expressões faciais registradas durante a visualização dos vídeos. Esses resultados destacam o potencial das expressões faciais como uma fonte de dados valiosa para compreender as emoções dos alunos em ambientes de aprendizagem online.

5.2.1.9 Lacunas de Pesquisa

Nos últimos anos, *transfer learning* emergiu como uma técnica poderosa em aprendizado de máquina, demonstrando sua eficácia em uma variedade de domínios, desde o reconhecimento de imagens até a tradução de idiomas. Na área da Educação, essa abordagem promete revolucionar a forma como os sistemas de aprendizagem são projetados e implementados, permitindo a transferência de conhecimento de uma tarefa ou domínio para outro, com o objetivo de melhorar a eficiência do ensino e da aprendizagem.

Apesar do crescente interesse e investimentos em *transfer learning* na Educação, identificou-se lacunas significativas nos estudos existentes que precisam ser abordadas para impulsionar ainda mais essa área promissora. Este trabalho propõe uma análise crítica das lacunas atuais em pesquisas relacionadas ao *transfer learning* na Educação, destacando áreas específicas que carecem de investigação aprofundada e discussão acadêmica.

Ao explorar essas lacunas, espera-se estimular novos questionamentos e direções de pesquisa que possam ampliar o entendimento sobre o potencial do *transfer learning* para transformar a Educação.

Ding *et al.* (2019) refletiu sobre a falta de disponibilidade de dados de *clickstream* em todos os MOOCs (ambientes de aprendizado online), representando uma lacuna significativa na pesquisa sobre *transfer learning* na Educação. Como observado pelos autores, essa limitação pode restringir a aplicabilidade da abordagem proposta, prejudicando a generalização dos resultados. Além disso, ao não levar em consideração informações contextuais latentes, como a fadiga ou motivação do aluno, a abordagem proposta pode estar negligenciando aspectos importantes que influenciam o comportamento de aprendizagem.

A falta de consideração dos objetivos e padrões de aprendizado pretendidos pelo professor também representa uma limitação, pois esses elementos desempenham um papel fundamental na eficácia do ensino e na condução do aprendizado do aluno. Por fim, a ausência de avaliação com usuários finais, como alunos, professores ou educadores, destaca a necessidade de experimentos adicionais para validar a eficácia da solução em um ambiente real de aprendizado online, destacando uma lacuna essencial que precisa ser preenchida para avançar no campo de *transfer learning* na Educação.

Huang e Guo (2021) em seu estudo ressaltaram questões fundamentais que permeiam a pesquisa em *transfer learning* na Educação. Em primeiro lugar, a dependência da qualidade dos dados de entrada, particularmente dos títulos e conteúdos dos artigos,

destaca a importância crítica de fontes de dados confiáveis e precisas para garantir a validade dos resultados obtidos. No entanto, ao reconhecerem que sua abordagem pode não ser adequada para avaliar a qualidade geral de um artigo, os autores apontaram para uma lacuna subjacente na capacidade dos modelos de *transfer learning* em discernir nuances complexas, além da mera relevância semântica entre o título e o conteúdo. Além disso, ao serem baseados em um modelo pré-treinado, os autores levantaram a questão da generalização e adaptabilidade da abordagem a tópicos específicos que podem não estar representados no modelo inicial.

Essa limitação aponta para a necessidade de desenvolver modelos mais flexíveis e adaptáveis que possam lidar com uma gama mais ampla de domínios e contextos educacionais. Por fim, a restrição da abordagem a estudos em inglês ressalta uma lacuna na representatividade e inclusão de múltiplos idiomas, destacando a importância de garantir que os avanços na pesquisa em *transfer learning* na Educação sejam aplicáveis globalmente e acessíveis a diversas comunidades linguísticas. Essas lacunas oferecem direcionamentos para futuras pesquisas e desenvolvimentos na área, visando uma abordagem mais robusta e abrangente para aplicação de *transfer learning* em contextos educacionais diversos.

Swamy *et al.* (2022) evidenciaram aspectos cruciais a serem considerados no contexto do estudo em questão. Em primeiro lugar, a limitação do conjunto de dados MOOC a uma única universidade na Europa, coletado há cinco anos, ressaltou a possibilidade de um viés inerente ao modelo devido à natureza dos dados subjacentes. Essa restrição levanta a necessidade premente de uma maior diversidade de fontes de dados, incluindo diferentes modalidades de cursos, como salas de aula invertidas e cursos mistos, a fim de avaliar adequadamente a transferibilidade independente da modalidade. Além disso, a dependência de recursos de comportamento artesanais extraídos dos fluxos de cliques brutos destaca uma lacuna na automação e eficiência do processo de extração de recursos.

A integração de modelos de recursos de comportamento latentes poderia mitigar essa limitação, oferecendo uma abordagem mais abrangente e integrada para a análise de dados, evitando a necessidade de extração manual de recursos. Essas lacunas identificadas apontam para áreas-chave que requerem atenção adicional e desenvolvimento futuro para aprimorar a robustez e a eficácia das abordagens de *transfer learning* em ambientes educacionais online.

Jensen *et al.* (2021) apontaram limitações importantes nas áreas de aprimoramento

e consideração para pesquisas futuras. Em primeiro lugar, a homogeneidade do conjunto de dados, composto por informações de apenas 16 professores de uma região geográfica similar, indicou uma possível limitação na generalização dos resultados para diferentes contextos linguísticos e culturais. Essa restrição ressalta a necessidade de uma maior diversidade de dados, abrangendo uma variedade de dialetos e origens geográficas, para garantir uma representação mais abrangente e precisa. Em segundo lugar, a consideração de um número limitado de modelos de aprendizagem automática revelou uma lacuna na exploração de toda a gama de abordagens disponíveis neste campo em constante evolução.

A inclusão de uma variedade mais ampla de modelos pode enriquecer a análise e fornecer uma compreensão mais abrangente sobre o uso dessas técnicas na análise de discurso em sala de aula. Por fim, a aplicabilidade específica dos resultados ao discurso em salas de aula de língua inglesa destacou a necessidade de expandir a investigação para outras disciplinas e áreas de estudo. Considerar métodos de *feedback* automatizado em diversas áreas disciplinares pode proporcionar uma compreensão mais completa do valor dessas ferramentas para os professores em diferentes contextos educacionais. Essas lacunas identificadas oferecem direções claras para pesquisas futuras, visando uma compreensão mais abrangente e aprimorada do uso de modelos de aprendizagem automática na Educação.

Em suma, as limitações destacadas pelos autores apontaram para áreas cruciais que merecem atenção e investigação adicional. Superar essas lacunas não apenas fortalecerá a validade e a generalização dos resultados, mas também promoverá avanços significativos no campo do processamento de linguagem natural e da análise de dados educacionais. Ao abordar questões como a diversidade dos conjuntos de dados, a variedade de modelos de aprendizagem automática e a aplicabilidade em diferentes contextos linguísticos e disciplinares, os pesquisadores podem contribuir para uma compreensão mais completa e eficaz do uso dessas tecnologias na Educação. Portanto, futuros estudos que se concentrem nessas áreas críticas têm o potencial impactar positivamente a prática educacional.

5.3 Avaliação de Qualidade dos Artigos Selecionados na RSL

A qualidade de um artigo científico é fundamental, pois reflete sua capacidade de atender a diversos padrões e critérios específicos. Segundo, Oliveira *et al.* (2021), esses critérios incluem a relevância do tema abordado, o rigor acadêmico na condução da pesquisa,

a clareza na exposição das ideias, a originalidade dos conteúdos apresentados, a robustez da metodologia empregada, a consistência das evidências fornecidas, a profundidade das discussões realizadas e a contribuição significativa para o corpo de conhecimento existente sobre o assunto investigado. Atender a esses critérios não só garante a credibilidade e a aceitação do trabalho pela comunidade acadêmica, como também assegura que o artigo possa servir como uma referência confiável para estudos futuros.

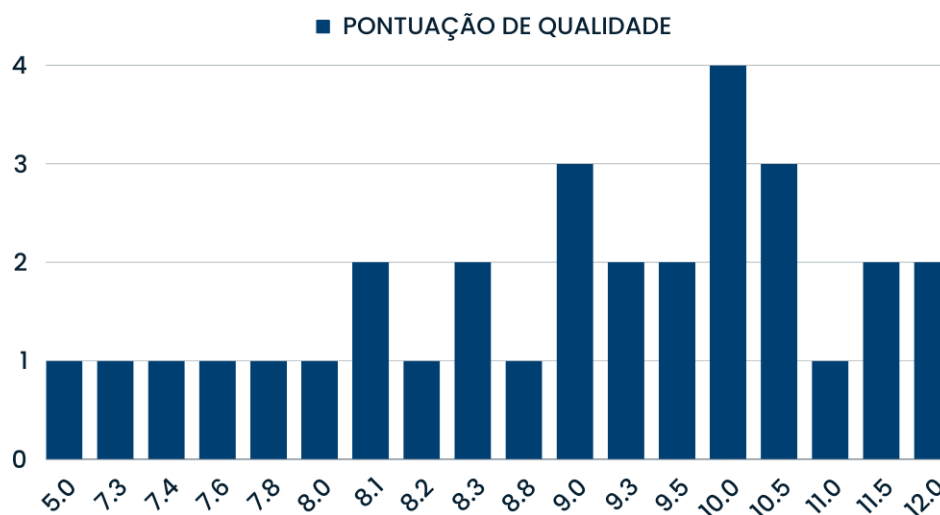
Avaliar a qualidade de um artigo depende do objetivo do trabalho em questão, seja para fins acadêmicos, científicos, jornalísticos ou outros. Greenhalgh (1997) destaca que a avaliação da qualidade de um artigo deve considerar não apenas a validade científica, mas também a relevância para a prática e o contexto em que o estudo será aplicado. Além disso, diferentes disciplinas podem ter critérios específicos para determinar a qualidade de um artigo. Por exemplo, enquanto a ciência médica pode valorizar ensaios clínicos randomizados e revisões sistemáticas, outras áreas podem priorizar metodologias qualitativas ou estudos de caso detalhados.

Em geral, a qualidade é uma medida subjetiva que pode variar de acordo com as expectativas e normas aceitas dentro de uma comunidade acadêmica ou profissional específica. De acordo com Smith (2020), essa subjetividade é influenciada por múltiplos fatores, incluindo os padrões metodológicos, a originalidade da pesquisa e o impacto esperado no campo de estudo.

Neste estudo, buscou-se demonstrar os índices de qualidade dos artigos utilizados na RSL, considerando as diferentes métricas ou critérios empregados para avaliar a excelência, relevância e confiabilidade do trabalho. Os índices de qualidade dos artigos são apresentados na Figura 7, oferecendo uma visão geral da qualidade dos 31 artigos analisados. Essa análise permitiu uma avaliação mais completa do corpus de estudos selecionados, destacando os trabalhos que se destacaram pela sua qualidade e contribuição para o avanço do conhecimento na área de estudo.

A pontuação dos artigos variou de 5.0 a 12.0, utilizando uma escala *4-Likert*. Essa escala, conhecida como Escala Likert em homenagem ao psicólogo Rensis Likert, é amplamente utilizada em pesquisas e questionários para avaliar atitudes, opiniões ou sentimentos das pessoas em relação a determinados itens. Segundo Likert (1932), sua aplicação é difundida em estudos sociais, psicológicos e de opinião, sendo uma ferramenta fundamental para a mensuração de variáveis subjetivas em pesquisa. Neste trabalho, a

Figura 7 – Visão geral da pontuação da avaliação de qualidade dos 31 artigos.



Fonte: Autoria Própria (2024).

escala foi ajustada para variar de 0.0 a 12.0, empregando métricas de validação e controle randomizado para conduzir essa análise.

Nesta pesquisa, foram estabelecidas categorias de qualidade fundamentadas em critérios amplamente reconhecidos na comunidade acadêmica para avaliar a robustez dos estudos. Segundo Grant e Booth (2009), esses critérios são essenciais para a avaliação objetiva da qualidade de pesquisa. A seguir, são descritas as características de cada categoria com base nos seguintes princípios:

1. **BAIXA:** a qualidade do estudo é considerada baixa quando possui uma metodologia inadequada, mal explicada ou que não permite uma análise confiável dos resultados, o que pode comprometer a qualidade do artigo;
2. **MÉDIA:** a qualidade do estudo é considerada média quando possui uma metodologia de pesquisa clara, bem explicada e adequada para responder às perguntas da pesquisa, mas pode haver espaço para melhorias ou aprimoramentos;
3. **ALTA:** a qualidade do estudo é considerada alta quando possui uma metodologia de pesquisa robusta, claramente definida e adequadamente aplicada, demonstrando uma abordagem científica sólida;
4. **MUITO ALTA:** a qualidade do estudo é considerada muito alta quando faz uma

contribuição excepcional e significativa para o conhecimento existente, introduzindo novos conceitos, teorias ou descobertas que têm um impacto substancial em seu campo de estudo.

Na jornada da pesquisa acadêmica, a qualidade dos estudos desempenha um papel fundamental. Compreender como essa qualidade é avaliada pode revelar informações importantes sobre a confiabilidade e a relevância de uma publicação. Dessa forma, ao selecionar os parâmetros no *software Parsifal* dedicados à análise de qualidade (Concordo totalmente, Concordo, Indiferente, Discordo e Discordo totalmente), o sistema equiparava essas escolhas com seus valores respectivos na escala de *4-Likert*. Por meio do somatório das notas de todos os critérios, os artigos foram então classificados em quatro faixas de qualidade, conforme apresentado na Tabela 6.

Tabela 6 – Pontuação para cada nível de qualidade.

BAIXA	MÉDIA	ALTA	MUITO ALTA
0 <= 4,9	5 <=7,9	8 <=9,9	10 <=12

De acordo com os dados obtidos das questões da Tabela 3, observou-se que o estudo de Gillani *et al.* (2023) obteve uma pontuação de 5.0 devido à falta de informações sobre o método de pesquisa e o uso de *transfer learning* na Educação. Além disso, o estudo de Ordóñez *et al.* (2002) apresentou a pontuação de 7.3, outro obteve Bassok e Holyoak (1989) 7.4, um terceiro Fan e Alwan (2022) 7.6, e um quarto atingiu Lindgren *et al.* (2022) 7.8. Essas pontuações refletem a escassez de detalhes na coleta de dados, falta de clareza nos resultados e pouca discussão sobre a relevância para a área da Educação.

Da mesma forma, um estudo Sharmila *et al.* (2023) obteve a pontuação de 8.0, enquanto dois estudos Karimah e Hasegawa (2022) e Vachev *et al.* (2021) alcançaram 8.1. Um estudo recebeu a pontuação de 8.2 Teran-Quezada *et al.* (2022), enquanto outros dois estudos Wang e Shi (2021) e Hardalov *et al.* (2020) foram avaliados com 8.3. Por fim, um estudo alcançou a pontuação de 8.8 Bian *et al.* (2019). Esses estudos reconheceram suas limitações, discutiram aspectos sobre a Educação e a técnica de *transfer learning*, no entanto, faltou mais detalhamento em seus resultados e métodos. Além disso, os estudos Jensen *et al.* (2021), Delaney e Bhatia (2021) e Benedetto *et al.* (2023) receberam a pontuação 9.0. Dois estudos obtiveram a pontuação 9.3: Alqahtani *et al.* (2023) e Hu *et al.*

(2022), enquanto outros dois alcançaram 9.5: Lehrer e Littlefield (1993) e Kahawanugoda *et al.* (2022). Quatro estudos apresentaram a pontuação 10.0: Xiao *et al.* (2022), Kunhoth *et al.* (2023), Ding *et al.* (2019) e Ikram *et al.* (2023). Três estudos receberam 10.5: Perseghin e Foresti (2023), Rosli *et al.* (2021) e Huang e Guo (2021).

Adicionalmente, um estudo obteve a pontuação 11.0: Lu (2022), enquanto dois estudos atingiram 11.5: Mor e Dardeck (2021) e Swamy *et al.* (2022). Por fim, dois estudos receberam a pontuação 12.0: Bonthu *et al.* (2023) e Yang *et al.* (2021). Essas avaliações refletem um nível elevado de qualidade nos estudos analisados. É relevante destacar que o termo "qualidade", neste contexto, está relacionado com a capacidade de também responder às perguntas de pesquisa definidas para este estudo.

A análise da qualidade dos estudos selecionados revelou uma variedade de pontuações, refletindo diferentes níveis de rigor metodológico, clareza na apresentação dos resultados e contribuição para o conhecimento existente. Enquanto alguns estudos apresentaram deficiências em áreas como metodologia de pesquisa e discussão dos resultados, outros se destacaram pela profundidade de sua abordagem, reconhecimento de limitações e relevância para a área da Educação e *transfer learning*.

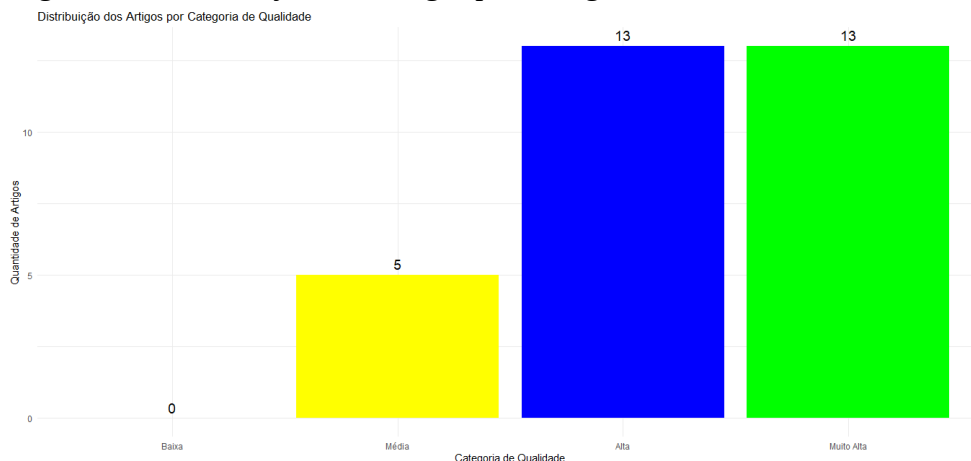
Essa diversidade de pontuações ressaltou a importância da avaliação criteriosa dos estudos durante o processo de execução da RSL. A adoção de uma abordagem rigorosa na seleção e avaliação dos estudos contribuiu para a confiabilidade e validade dos resultados da pesquisa, permitindo que fossem extraídas conclusões significativas e robustas para uso posterior pela comunidade científica.

Para uma visualização mais clara da qualidade dos estudos analisados, a Figura 8 apresenta a distribuição dos artigos em cada categoria de qualidade. Este gráfico ilustra a quantidade de artigos em cada faixa de qualidade, facilitando a compreensão da variação entre as categorias.

A ausência de artigos classificados como de qualidade baixa sugere que a maioria dos estudos analisados possui uma qualidade mínima aceitável, indicando uma base sólida de pesquisa em Transfer Learning e suas aplicações na educação.

Na categoria média, estão presentes 5 estudos com abordagens intermediárias, que são valiosos por sua contribuição ao entendimento e aplicação do Transfer Learning. Contudo, suas pontuações sugerem que ainda há espaço para melhorias na metodologia, na profundidade da análise ou na aplicabilidade dos resultados. A diversidade de tópicos

Figura 8 – Distribuição dos Artigos por Categoria de Qualidade.



Fonte: Autoria Própria (2024).

abordados reflete a ampla aplicação do Transfer Learning em diversas áreas educacionais.

Em contraste, 13 artigos são classificados como de qualidade alta, destacando-se por suas contribuições significativas e aplicação prática robusta. Esses estudos abordam questões críticas, como a detecção precoce de dislexia e disgrafia, o reconhecimento de gestos e emoções em contextos educacionais e a personalização do aprendizado online. Eles evidenciam a eficácia do Transfer Learning em aprimorar a precisão e a relevância das ferramentas educacionais, além de abrir caminho para novas pesquisas e inovações no campo.

Na categoria muito alta, 13 artigos representam o ápice da pesquisa em Transfer Learning aplicada à educação. Com pontuações elevadas, esses estudos refletem excelência em inovação, metodologia e impacto. Exemplos incluem o desenvolvimento de sistemas para detecção de violência em escolas e a previsão de sucesso em MOOCs, que demonstram o potencial transformador do Transfer Learning. Esses trabalhos não só oferecem soluções práticas para desafios educacionais contemporâneos, mas também estabelecem novas direções para futuras pesquisas.

A distribuição dos artigos revela um campo de pesquisa dinâmico e diversificado, com muitos estudos alcançando alta qualidade. A ausência de artigos na categoria baixa é encorajadora, sugerindo que a pesquisa em Transfer Learning está avançando com rigor e relevância. No entanto, a variação entre os níveis de profundidade e aplicação prática nas categorias média, alta e muito alta indica áreas onde ainda há espaço para melhorias metodológicas e maior aplicabilidade prática.

5.4 Visão Geral dos Resultados da Meta-análise

A meta-análise agrupa vários resultados de estudos em um único resultado, buscando expressar resultados semelhantes. Esta técnica permite combinar dados de múltiplas investigações, aumentando o poder estatístico e a generalização dos achados. Conforme Borenstein *et al.* (2009), a meta-análise sintetiza a literatura existente e fornece uma visão integrada sobre um campo de estudo. Ela combina os resultados de estudos independentes para obter uma estimativa geral mais precisa dos efeitos ou relações investigados.

Com a aplicação da meta-análise, os pesquisadores podem transcender as limitações dos estudos individuais, aproveitando a variedade de dados disponíveis para obter uma estimativa geral mais robusta e confiável dos efeitos ou relações em questão. Além disso, a meta-análise fornece uma estimativa quantitativa precisa do efeito médio ou da relação entre variáveis, oferecendo uma medida objetiva que sintetiza a evidência acumulada de forma abrangente e precisa.

Ao combinar e analisar os resultados de estudos independentes, a meta-análise amplia o entendimento, permitindo *insights* mais profundos e conclusões mais sólidas. Ela não apenas ajuda a identificar padrões e tendências comuns, mas também a detectar variações e discrepâncias que podem oferecer informações valiosas para futuras investigações. Em essência, a meta-análise é uma ferramenta essencial para os cientistas, capacitando-os a explorar as complexidades do conhecimento científico com maior clareza, precisão e confiabilidade.

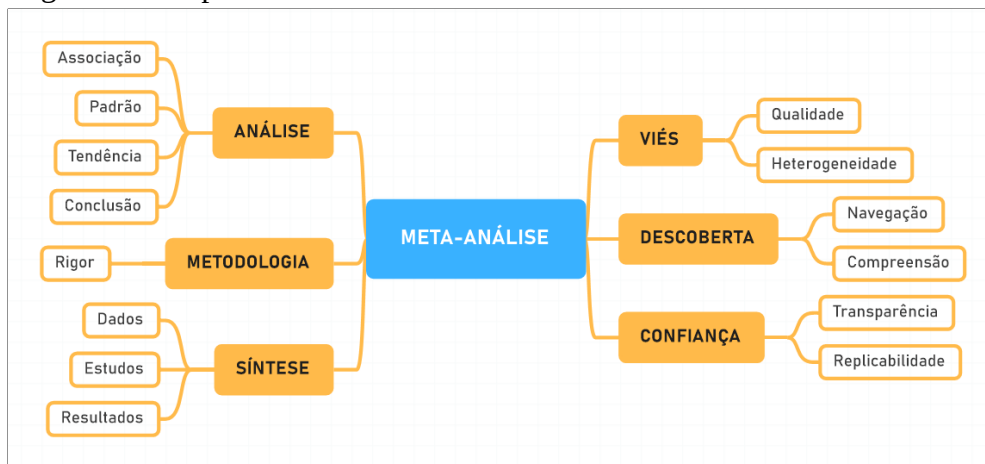
5.4.1 Resultado do Estudo sobre Meta-análise

Um dos aspectos mais peculiares da meta-análise é sua capacidade de transcender as limitações de estudos individuais. Mesmo quando os resultados de um estudo específico podem ser obscuros ou contraditórios, a meta-análise pode revelar a verdade subjacente, ampliando o entendimento além das idiosincrasias de qualquer pesquisa isolada. Segundo Hunter e Schmidt (2004), essa abordagem metodológica é essencial para obter uma visão mais abrangente e precisa dos efeitos ou relações investigadas.

Além disso, a meta-análise lembra da importância da transparência e da replicabilidade na pesquisa científica. Ela exige rigor metodológico e uma abordagem sistemática para garantir conclusões que sejam sólidas e confiáveis. Conforme ressaltam Higgins e

Green (2011), essas práticas são cruciais para assegurar que os resultados obtidos sejam válidos e possam ser reproduzidos por outros pesquisadores.

Figura 9 – Mapa Mental sobre o estudo sobre Meta-análise.



Fonte: Autoria Própria (2024).

As aplicações da meta-análise em diversos campos são exploradas, enfatizando sua utilidade na avaliação da eficácia de intervenções, políticas e métodos. Ao reconhecer as vantagens da meta-análise, como o aumento do poder estatístico e a capacidade de generalização, também são discutidos os desafios, incluindo a heterogeneidade entre estudos e o viés de publicação. Em resumo, o mapa mental sobre meta-análise é uma ferramenta valiosa para pesquisadores, estudantes e profissionais interessados em compreender e aplicar esse método fundamental de síntese de evidências.

5.4.2 Resultado da Identificação de Softwares

A condução de meta-análises desempenha um papel essencial na síntese de evidências provenientes de estudos. Para realizar essa tarefa de maneira eficaz, o uso de *software* estatístico especializado é fundamental. Diversos *softwares* são amplamente utilizados para a condução de meta-análises, incluindo o RevMan, Comprehensive Meta-Analysis (CMA), a Linguagem e Ambiente *R*, e o Stata. Dentre essas opções, foi utilizado o *R*. A escolha pelo *R* deve-se a diversas vantagens em relação às demais ferramentas. O *R* destaca-se por sua flexibilidade e capacidade de personalização, permitindo ajustes específicos conforme os requisitos metodológicos de cada estudo.

Além disso, sua extensa coleção de pacotes desenvolvidos pela comunidade oferece uma variedade de ferramentas estatísticas e de visualização projetadas especialmente para

meta-análises. Outro ponto positivo é a comunidade ativa de usuários e desenvolvedores, que contribui constantemente com novos pacotes e atualizações, mantendo o *R* na vanguarda das metodologias estatísticas. Esses fatores embasaram a escolha e justificam a preferência pelo *R* em comparação com outras ferramentas identificadas.

A flexibilidade do *R* permite a personalização de análises conforme as necessidades específicas de cada estudo, facilitando a implementação de modelos estatísticos avançados e a visualização de resultados. Portanto, o uso do *R* não só melhora a precisão e a eficiência das meta-análises, mas também promove a inovação e a adaptação às novas exigências da pesquisa científica.

O software *R* é amplamente usado por sua flexibilidade e capacidade de personalização. Os pesquisadores podem ajustar facilmente os procedimentos de análise de acordo com seus estudos e requisitos metodológicos específicos. Além disso, o *R* é valorizado por sua extensa coleção de pacotes desenvolvidos pela comunidade, oferecendo uma variedade de ferramentas estatísticas e de visualização projetadas especialmente para meta-análises.

O uso do *R* para conduzir meta-análises oferece diversas vantagens significativas, além de algumas desvantagens que devem ser cuidadosamente avaliadas. Na Tabela 7 são sintetizadas as vantagens e desvantagens mapeadas do uso do software *R* para conduzir meta-análises.

As vantagens e desvantagens do uso do *software R* para a condução de meta-análises foram mapeadas para fornecer uma visão clara e abrangente das suas capacidades e limitações. Entre as principais vantagens do uso do *R* para conduzir meta-análises, destacam-se a vasta gama de pacotes e bibliotecas disponíveis através do *Comprehensive R Archive Network (CRAN)*, que proporciona ferramentas especializadas para análises estatísticas e visualizações. O *R* é amplamente reconhecido por suas capacidades estatísticas avançadas e por suas poderosas ferramentas de visualização, como o *ggplot2*, que é um dos pacotes mais populares e poderosos para visualização de dados na linguagem de programação *R*, que permitem criar gráficos detalhados e personalizados. Além disso, a grande e ativa comunidade *R* oferece uma abundância de recursos e suporte, e sua natureza de código aberto permite personalização e integração com outras linguagens de programação.

Por outro lado, o *R* apresenta algumas desvantagens, como uma curva de aprendizado acentuada para aqueles novos em programação e estatística. O gerenciamento de memória e o desempenho com grandes conjuntos de dados também podem ser desafiadores,

Tabela 7 – Vantagens e Desvantagens do uso do R para conduzir Meta-análises

Vantagens	Desvantagens
Pacotes e Extensibilidade: O R possui uma vasta gama de pacotes e bibliotecas desenvolvidas pela comunidade para realizar diversas análises estatísticas, visualizações e manipulação de dados. O CRAN (Comprehensive R Archive Network) é um repositório centralizado que oferece acesso a milhares de pacotes.	Curva de Aprendizado: Para quem não está familiarizado com programação, o R pode ter uma curva de aprendizado acentuada, especialmente para usuários novos em estatística e programação.
Capacidade Estatística Avançada: O R é conhecido por suas capacidades estatísticas robustas, incluindo testes estatísticos, modelos de regressão, análises de séries temporais, e técnicas avançadas de análise multivariada.	Desempenho com Grandes Conjuntos de Dados: O R pode enfrentar dificuldades de desempenho ao lidar com grandes volumes de dados, embora isso possa ser mitigado com técnicas e pacotes específicos, como o <code>data.table</code> .
Visualização de Dados: O R oferece poderosas ferramentas de visualização, como <code>ggplot2</code> , que permitem criar gráficos detalhados e personalizados.	Gerenciamento de Memória: O gerenciamento de memória no R pode ser menos eficiente em comparação com outras linguagens, como Python, especialmente quando se trabalha com grandes conjuntos de dados.
Comunidade Ativa: A comunidade R é grande e ativa, o que significa que há uma abundância de recursos, tutoriais e fóruns de suporte para ajudar os usuários.	Interface do Usuário: Embora o RStudio tenha melhorado significativamente a interface do usuário, o R, em sua forma padrão, não é tão intuitivo quanto algumas ferramentas baseadas em menus.
Código Aberto: Sendo uma ferramenta de código aberto, o R é gratuito para usar e modificar, o que permite personalização e compartilhamento de ferramentas e métodos.	Documentação Variada: A qualidade da documentação pode variar entre pacotes e funções, o que pode dificultar a busca por informações precisas e confiáveis.
Integração com Outras Linguagens: O R pode ser integrado com outras linguagens de programação como C++, Python e Java, facilitando a integração com outras ferramentas e tecnologias.	Atualizações e Compatibilidade: Atualizações frequentes de pacotes podem ocasionalmente causar problemas de compatibilidade, exigindo ajustes e manutenção contínuos do código.

embora existam técnicas e pacotes para mitigar esses problemas. Além disso, a interface padrão do R pode não ser tão intuitiva quanto algumas ferramentas baseadas em menus, e a qualidade da documentação dos pacotes pode variar, o que pode dificultar a busca por informações precisas.

Em resumo, o software R desempenha um papel fundamental na pesquisa e prática científica atual, oferecendo uma gama incomparável de funcionalidades e oportunidades para explorar e compreender os dados de forma mais profunda e significativa.

5.4.3 Resultado da Definição de Modelos

O modelo de efeitos aleatórios é uma abordagem estatística essencial em revisões sistemáticas e meta-análises, especialmente em estudos de *transfer learning* na Educação. Segundo DerSimonian e Kacker (2007), esse modelo é crucial para capturar a variabilidade entre os estudos incluídos. Ele é escolhido devido à diversidade nas características, métodos utilizados e outros fatores que podem influenciar os resultados dos estudos. Essa abordagem permite considerar a heterogeneidade entre os estudos, sendo fundamental devido às variações nas estimativas de efeitos do tratamento. Ao contrário do modelo de efeitos fixos, que assume um único efeito verdadeiro compartilhado por todos os estudos, o modelo de efeitos aleatórios captura essa heterogeneidade e a variabilidade nos dados dos ensaios clínicos analisados na meta-análise.

Esse modelo é amplamente empregado em meta-análises para acomodar a variabilidade significativa observada entre os estudos incluídos. Higgins *et al.* (2003) discutem como este modelo estatístico é essencial para lidar não apenas com a heterogeneidade dentro de cada estudo, mas também para capturar as diferenças substanciais entre os estudos em termos de características populacionais, metodológicas ou outras fontes de variabilidade. Ao permitir que os pesos dos estudos variem de acordo com a precisão de cada estimativa, o modelo de efeito aleatório melhora a robustez das conclusões da meta-análise, fornecendo uma estimativa agregada mais precisa e representativa dos efeitos de interesse.

Em uma revisão sistemática sobre *transfer learning* na Educação, o modelo aleatório permite analisar os efeitos de intervenções ou técnicas de *transfer learning* em diferentes contextos educacionais. Reconhecendo que os estudos individuais variam em características dos participantes, métodos de intervenção, medidas de resultado e outros fatores. Assim, o modelo aleatório pode atribuir pesos a cada estudo com base em sua precisão e contribuição para a estimativa geral do efeito.

Ao incorporar o modelo aleatório, os pesquisadores podem obter uma estimativa mais precisa e abrangente do efeito do *transfer learning*, considerando a heterogeneidade entre os estudos. Isso permite uma análise robusta e generalizável dos resultados, oferecendo informações valiosas para informar práticas educacionais e futuras pesquisas.

A escolha de modelos aleatórios em meta-análises é motivada por várias razões. Segundo Higgins e Green (2011) e Borenstein *et al.* (2009), eles são capazes de lidar com

a heterogeneidade entre os estudos, sintetizando evidências de estudos com diferentes populações, intervenções e desenhos de estudo. Além disso, de acordo com Deeks *et al.* (2019) e IntHout *et al.* (2016), atribuem pesos aos estudos com base em sua precisão e contribuição, possibilitando uma análise equilibrada e robusta. Outras vantagens, segundo Riley *et al.* (2011) e Viechtbauer (2010), incluem a capacidade de generalização para diferentes contextos e populações, a flexibilidade estatística e a robustez a vieses desconhecidos. Essas características, de acordo com os autores citados, tornam os modelos aleatórios uma escolha preferencial para meta-análises, proporcionando estimativas mais precisas e confiáveis dos efeitos de intervenções ou tratamentos.

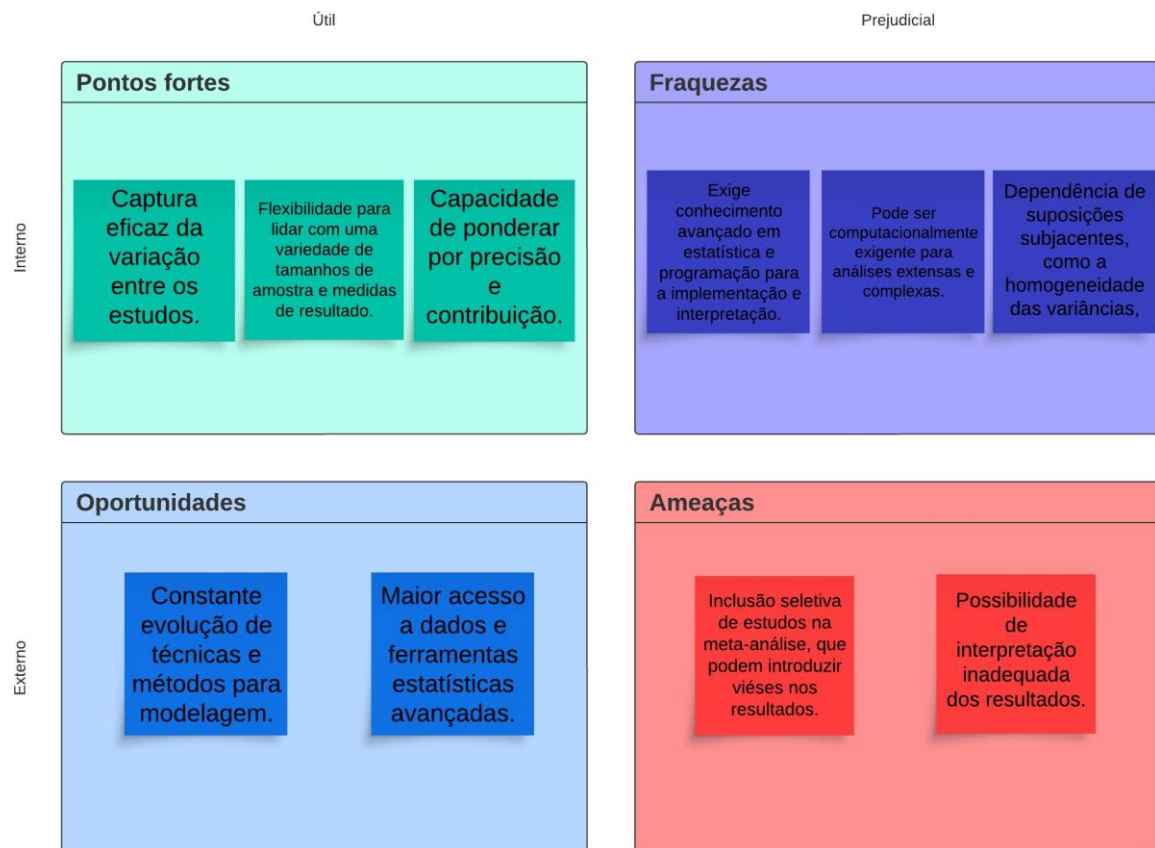
A análise *Strengths, Weaknesses, Opportunities, Threats* (SWOT) é comumente utilizada em diversos contextos, incluindo negócios, planejamento estratégico, marketing, gestão de projetos e pesquisa acadêmica. Na pesquisa acadêmica, a análise SWOT oferece uma estrutura para avaliar teorias e metodologias, destacando suas forças e fraquezas e identificando áreas de oportunidade e possíveis ameaças. Esse enfoque permite uma avaliação crítica e detalhada das opções metodológicas disponíveis, facilitando a identificação de abordagens que melhor se adaptem às necessidades específicas de cada estudo.

Foi realizada uma análise SWOT (ver Figura 10), para avaliar a aplicação do modelo aleatório, seguindo a metodologia descrita por Campbell (2023). Devido à heterogeneidade dos estudos, a análise SWOT focou exclusivamente no modelo aleatório, ajudando a identificar seus pontos fortes, fracos, oportunidades e ameaças para garantir uma abordagem adequada à variabilidade dos dados.

Com isso, é possível oferecer uma visão abrangente e estruturada dos principais aspectos relacionados à escolha do método aleatório em uma meta-análise sobre *transfer learning* na Educação. A análise SWOT contribui para garantir uma abordagem mais robusta e informada para lidar com a variabilidade dos dados, permitindo que pesquisadores tomem decisões fundamentadas e aprimorem a qualidade e a relevância de suas conclusões.

Ao identificar os pontos fortes, pontos fracos, oportunidades e ameaças associadas a essa decisão, foi possível avaliar de forma sistemática os diferentes aspectos que podem influenciar o sucesso da meta-análise. Isso proporciona uma base sólida para tomada de decisão informada, permitindo a identificação de áreas que podem ser maximizadas, bem como áreas que exigem atenção especial ou mitigação de riscos. Em última análise, a análise SWOT serve como uma ferramenta estratégica para orientar o planejamento e a

Figura 10 – Análise SWOT da escolha do Método Aleatório em uma Meta-análise sobre *transfer learning* na Educação.



Fonte: Autoria Própria (2024).

execução da meta-análise, ajudando a garantir que as decisões sejam fundamentadas em uma compreensão completa e equilibrada dos fatores envolvidos.

A análise SWOT apresentada na Figura 10 destaca os principais aspectos relacionados à escolha do método aleatório. Os pontos fortes incluem a flexibilidade estatística para lidar com diferentes desenhos de estudo, enquanto os pontos fracos destacam a complexidade estatística e a dependência de suposições subjacentes. As oportunidades identificadas incluem a constante evolução de técnicas e métodos estatísticos, enquanto as ameaças destacam os riscos associados a vieses de publicação e interpretação inadequada dos resultados. Essa análise fornece uma visão abrangente dos fatores que devem ser considerados ao tomar decisões sobre o uso de modelos aleatórios em meta-análises sobre *transfer learning* na Educação.

5.4.4 Resultado da Avaliação da Heterogeneidade dos Estudos

Nesta seção são apresentados os resultados da avaliação da heterogeneidade dos grupos envolvidos no estudo de *transfer learning* na Educação. Os resultados apresentados foram derivados de uma análise dos 31 estudos identificados na Revisão Sistemática da Literatura (RSL). Destes, apenas 25 estudos foram utilizados devido à falta de disponibilidade de dados nos demais. Para compreender a diversidade de desempenho entre os grupos, foi realizada uma análise detalhada, utilizando gráficos de comparação de desempenho e gráficos de floresta.

O gráfico de comparação de desempenho entre diferentes grupos de estudos oferece uma visão abrangente das diferenças quantitativas nos resultados obtidos. Cada grupo é representado por uma barra, permitindo uma comparação visual imediata. Esse tipo de representação é essencial para identificar disparidades significativas de desempenho entre os grupos e pode revelar informações valiosas sobre os fatores que influenciam os resultados.

Além disso, utilizou-se um gráfico de floresta para explorar a contribuição relativa de diferentes variáveis na heterogeneidade observada nos resultados. Segundo o estudo de Higgins e Thompson (2002), esse método permite identificar quais variáveis têm o maior impacto na variação dos resultados entre os grupos. Cada nó no gráfico de floresta representa uma variável, enquanto as ramificações indicam sua influência na heterogeneidade. Essa abordagem ajuda a identificar padrões complexos e a compreender melhor os determinantes da diversidade de desempenho.

Essas análises combinadas proporcionam uma compreensão abrangente da heterogeneidade presente no estudo de *transfer learning* na Educação, destacando tanto as diferenças quantitativas quanto os fatores subjacentes que as impulsionam.

5.4.4.1 Comparação de Desempenho entre Grupos

Nesta seção é conduzida uma análise detalhada da comparação de desempenho entre diferentes grupos, com base em dados coletados de uma série de estudos no contexto educacional utilizando *Transfer Learning*. O objetivo foi identificar padrões e tendências nos resultados de desempenho e explorar possíveis correlações entre variáveis, como o número de estudos e o desempenho médio. Ao longo desta análise, examinou-se como

cada grupo se saiu nas respectivas métricas de desempenho utilizadas no estudo e discuti-se as implicações desses resultados.

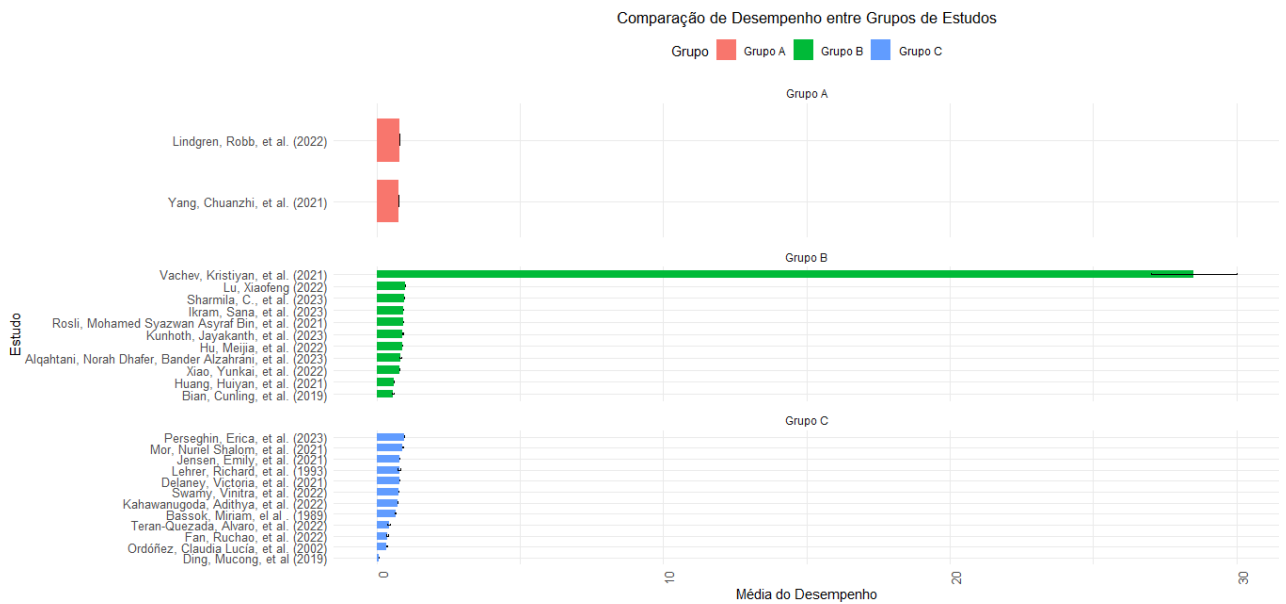
Na análise comparativa de desempenho entre grupos de estudos, foram selecionados 25 estudos que forneceram informações completas ou parciais de média de desempenho, desvio padrão e tamanho da amostra. A escolha dos estudos foi guiada pela disponibilidade desses dados essenciais para uma análise robusta. É importante notar que, enquanto este estudo se baseou em 25 estudos, a Referência Sistemática da Literatura (RSL) incluiu um total de 31 estudos. A diferença ocorreu devido à exclusão de estudos que não apresentaram informações suficientes para comparação direta de desempenho. Dentro dos grupos identificados - Grupo A, Grupo B e Grupo C - observou-se que o Grupo A apresentou a melhor média de desempenho, seguido pelo Grupo B e, por último, pelo Grupo C. Esses resultados sugerem que as abordagens ou contextos representados pelo Grupo A podem ter um desempenho superior com base nos dados analisados.

O Grupo A foi avaliado utilizando métricas de desempenho gerais, que incluem testes pré e pós, teste de transferência atrasado, avaliação de engajamento e análise de vídeo. No Grupo B foram aplicadas as métricas de Precisão e *Recall*, como precisão, *recall*, *AP*, *mAP@.5* e *mAP@.5:0.95*, com base nos resultados de testes específicos. Por fim, o Grupo C foi avaliado usando outras métricas de avaliação, como Área sob a Curva, *Word Error Rate (WER%)*, *Balanced Accuracy (BAC)*, precisão de treinamento e validação, *AUROC*, *Gwet's AC*, entre outras.

Inicialmente, foram examinadas as médias de desempenho dos Grupos A, B e C, conforme apresentado na Figura 11. A análise destaca as diferenças observadas entre eles, proporcionando uma visão clara de como cada grupo se comporta em termos de desempenho. Esses resultados permitem identificar as abordagens ou contextos que possivelmente contribuem para um desempenho superior. O Grupo B apresentou a melhor média de desempenho entre os grupos, seguido pelo Grupo A, que também mostrou um desempenho elevado, mas com maior variabilidade. Por último, o Grupo C apresentou as médias de desempenho mais baixas.

Em seguida, foi discutido possíveis interpretações para essas discrepâncias, considerando fatores como o tamanho do grupo e a qualidade dos estudos incluídos. Por fim, são apresentadas algumas reflexões sobre a importância dessas análises para a compreensão mais ampla dos resultados e para orientar futuras pesquisas na área. Esta análise proporci-

Figura 11 – Gráfico de Comparação de Desempenho entre Grupos de Estudos.



Fonte: Autoria Própria (2024).

ona informações valiosas para o campo do *Transfer Learning* na Educação, ajudando a direcionar esforços e recursos de pesquisa de forma mais eficaz.

A análise do gráfico comparativo revela que o Grupo B apresentou as médias de desempenho mais altas, com intervalos de confiança relativamente estreitos, indicando maior precisão nas estimativas dos estudos desse grupo. O Grupo A também teve boas médias, mas com uma variabilidade mais alta nos resultados, indicada por intervalos de confiança mais largos. O Grupo C apresentou as médias de desempenho mais baixas em comparação com os outros grupos, e seus intervalos de confiança amplos sugerem uma menor precisão nas estimativas, refletindo um padrão de desempenho geralmente inferior. Essas observações destacam a importância de considerar não apenas as médias de desempenho, mas também a consistência e a variabilidade dos resultados para uma avaliação mais completa das diferentes abordagens ou contextos de estudo.

Portanto, através desta análise detalhada, observou-se uma notável variação no desempenho e no desvio padrão entre os estudos, refletindo uma ampla diversidade nos resultados. As médias de desempenho exibem uma variação significativa, com alguns estudos registrando médias próximas de zero, enquanto outros se aproximam de 1. Isso aponta para uma disparidade considerável nos resultados obtidos entre os estudos, sugerindo a presença de fatores variáveis que influenciam o desempenho do *Transfer Learning* na Educação.

Na Tabela 8 é apresentada uma comparação detalhada do desempenho entre diferen-

tes grupos de estudos, cada linha representando um estudo específico dentro desses grupos. Os dados incluem a média, desvio padrão e tamanho da amostra de cada estudo. Os grupos foram agrupados em três categorias com base nas métricas de avaliação utilizadas (Grupo A, Grupo B e Grupo C). O grupo A para métricas gerais de desempenho, Grupo B para métricas de Precisão e *Recall*, e Grupo C para outras métricas como Área sob a Curva e *Word Error Rate (WER)*.

Tabela 8 – Comparação de Desempenho entre Grupos de Estudos

Grupo	Estudo	Média	Desvio Padrão	Tamanho da Amostra
Grupo A	Lindgren, Robb, et al. (2022)	0.800	0.050	80
Grupo A	Yang, Chuanzhi, et al. (2021)	0.777	0.007	210,488
Grupo B	Vachev, Kristiyan, et al. (2021)	28.51	7.57	100.000
Grupo B	Hu, Meijia, et al. (2022)	0.910	0.028	4,550
Grupo B	Ikram, Sana, et al. (2023)	0.930	0.000	3,000
Grupo B	Huang, Huiyan, et al. (2021)	0.600	0.008	10,000
Grupo B	Kunhoth, Jayakanth, et al. (2023)	0.910	0.034	45
Grupo B	Bian, Cunling, et al. (2019)	0.577	0.158	82
Grupo B	Rosli, Mohamed Syazwan Asyraf Bin, et al. (2021)	0.920	0.033	138,500
Grupo B	Xiao, Yunkai, et al. (2022)	0.810	0.061	7,418
Grupo B	Lu, Xiaofeng (2022)	0.980	0.007	60
Grupo B	Alqahtani, Norah Dhafer, Bander Alzah-rani, et al. (2023)	0.835	0.085	30
Grupo B	Sharmila, C., et al. (2023)	0.950	0.029	3,000
Grupo C	Ding, Mucong, et al (2019)	0.080	0.000	52,562
Grupo C	Kahawanugoda, Adithya, et al. (2022)	0.732	0.237	1,450
Grupo C	Swamy, Vinitra, et al. (2022)	0.750	0.000	145,714
Grupo C	Delaney, Victoria, et al. (2021)	0.786	0.004	2,000
Grupo C	Jensen, Emily, et al. (2021)	0.812	0.029	16,977
Grupo C	Fan, Ruchao, et al. (2022)	0.366	0.265	549
Grupo C	Teran-Quezada, Alvaro, et al. (2022)	0.420	0.326	195
Grupo C	Mor, Nuriel Shalom, et al. (2021)	0.910	0.019	50
Grupo C	Perseghin, Erica, et al. (2023)	0.950	0.000	1,800
Grupo C	Bassok, Miriam, el al . (1989)	0.653	0.189	164
Grupo C	Ordóñez, Claudia Lucía, et al. (2002)	0.345	0.105	88
Grupo C	Lehrer, Richard, et al. (1993)	0.790	0.151	47

A análise dos dados apresentados na Tabela 8 oferece uma visão detalhada do desempenho dos diferentes grupos de estudos. Os grupos foram avaliados com base em médias, desvios padrão e tamanhos de amostra, permitindo uma comparação abrangente entre eles.

O Grupo A demonstra um desempenho consistentemente elevado, com médias de 0.800 e 0.777, resultando em uma média geral de 0.7884. Os desvios padrão baixos (0.050 e 0.007) indicam uma menor variabilidade e maior precisão nas estimativas de desempenho. Esses resultados sugerem que os estudos incluídos no Grupo A possuem uma alta confiabilidade, evidenciando que seus métodos e resultados são bastante consistentes.

Em contraste, o Grupo B apresenta uma ampla variação nas médias de desempenho, que oscilam entre 0.08 e 28.51, com uma média geral de aproximadamente 3.178. Essa dispersão é acompanhada por desvios padrão significativamente variáveis (de 0.000 a 7.57), indicando uma considerável heterogeneidade entre os estudos do grupo. A alta variabilidade pode ser atribuída a diferenças metodológicas ou contextuais, sugerindo que o Grupo B inclui abordagens diversas que impactam os resultados de desempenho de maneiras distintas.

O Grupo C mostra um desempenho geral inferior, com médias que variam de 0.345 a 0.950 e uma média geral de 0.618. Embora o desempenho médio seja mais baixo em comparação com os Grupos A e B, o Grupo C exibe uma consistência relativamente maior em seus resultados, com desvios padrão moderados. Isso sugere que, apesar das médias serem inferiores, os estudos do Grupo C são relativamente homogêneos em termos de desempenho, indicando que podem estar enfrentando desafios ou características comuns que afetam os resultados.

Em resumo, o Grupo A se destaca com um desempenho superior e mais estável, com uma média geral de 0.7884 e baixos desvios padrão, indicando alta precisão nas estimativas de desempenho. No entanto, o Grupo B apresenta a maior média geral de 3.178, embora com uma variabilidade considerável, o que pode refletir uma diversidade de abordagens e contextos entre os estudos desse grupo. Por outro lado, o Grupo C, apesar de ter uma média geral inferior de 0.618, demonstra uma consistência interna relativamente alta em seus resultados. Essas observações ressaltam a importância de considerar tanto a média quanto a variabilidade dos dados ao avaliar o desempenho dos estudos, além dos fatores contextuais que podem influenciar essas diferenças.

Adicionalmente, os intervalos de confiança fornecem uma medida crucial da incerteza associada às estimativas de desempenho. Estudos que apresentam intervalos de confiança mais estreitos indicam resultados mais precisos e uma maior confiança nas estimativas de desempenho, enquanto intervalos mais amplos sugerem uma maior incerteza e uma necessidade de cautela na interpretação dos resultados.

Na Tabela 9 são apresentados os resultados dos estudos selecionados, organizados por grupo e acompanhados dos intervalos de confiança (IC). Cada estudo é identificado pelo seu título completo, seguido pelos valores correspondentes de IC inferior e superior. Esses intervalos são essenciais para avaliar a precisão das estimativas encontradas em cada

grupo de estudo, fornecendo uma visão detalhada da variabilidade dos resultados obtidos.

Tabela 9 – Estudos com Intervalo de Confiança

Grupo	Estudo	IC Inferior	IC Superior
Grupo A	Lindgren, Robb, et al. (2022)	0.789	0.811
Grupo A	Yang, Chuanzhi, et al. (2021)	0.776	0.779
Grupo B	Vachev, Kristiyan, et al. (2021)	28.46	28.56
Grupo B	Hu, Meijia, et al. (2022)	0.905	0.915
Grupo B	Ikram, Sana, et al. (2023)	0.930	0.930
Grupo B	Huang, Huiyan, et al. (2021)	0.599	0.601
Grupo B	Kunhoth, Jayakanth, et al. (2023)	0.896	0.924
Grupo B	Bian, Cunling, et al. (2019)	0.538	0.616
Grupo B	Rosli, Mohamed Syazwan Asyraf Bin, et al. (2021)	0.919	0.921
Grupo B	Xiao, Yunkai, et al. (2022)	0.806	0.814
Grupo B	Lu, Xiaofeng (2022)	0.977	0.983
Grupo B	Alqahtani, Norah Dhafer, Bander Alzah-rani, et al. (2023)	0.797	0.873
Grupo B	Sharmila, C., et al. (2023)	0.949	0.951
Grupo C	Ding, Mucong, et al (2019)	0.080	0.080
Grupo C	Kahawanugoda, Adithya, et al. (2022)	0.665	0.799
Grupo C	Swamy, Vinitra, et al. (2022)	0.750	0.750
Grupo C	Delaney, Victoria, et al. (2021)	0.784	0.788
Grupo C	Jensen, Emily, et al. (2021)	0.811	0.813
Grupo C	Fan, Ruchao, et al. (2022)	0.273	0.458
Grupo C	Teran-Quezada, Alvaro, et al. (2022)	0.237	0.603
Grupo C	Mor, Nuriel Shalom, et al. (2021)	0.872	0.948
Grupo C	Perseghin, Erica, et al. (2023)	0.950	0.950
Grupo C	Bassok, Miriam, et al. (1989)	0.527	0.779
Grupo C	Ordóñez, Claudia Lucía, et al. (2002)	0.139	0.551
Grupo C	Lehrer, Richard, et al. (1993)	0.641	0.939

A tabela 9 fornece uma visão detalhada dos intervalos de confiança (IC) para diferentes estudos categorizados em três grupos distintos: Grupo A, Grupo B e Grupo C. Essas análises ressaltam a importância de considerar tanto a média quanto a variabilidade dos dados ao avaliar o desempenho dos estudos, bem como os fatores contextuais que podem influenciar essas diferenças.

O Grupo A mostra intervalos de confiança relativamente estreitos, indicando uma alta precisão nas estimativas de desempenho dos estudos. Por exemplo, o estudo de Lindgren *et al.* (2022) tem um IC de 0.789 a 0.811, o que sugere uma média de desempenho bem definida e precisa. Similarmente, o estudo de Yang *et al.* (2021) tem um IC de 0.776 a 0.779, refletindo uma precisão similar. Esses resultados são consistentes com a média mais alta observada para o Grupo A, com valores próximos a 0.7884, destacando a estabilidade e a confiabilidade dos estudos neste grupo.

No Grupo B, os intervalos de confiança variam significativamente. O estudo de

Vachev *et al.* (2021) tem um IC extremamente largo, de 68.079 a 81.721, refletindo uma grande variabilidade e possíveis questões de precisão na estimativa. Em contraste, estudos como o de Lu (2022) mostram um IC mais estreito de 0.977 a 0.983, indicando alta precisão e um desempenho consistentemente alto. A ampla gama de ICs no Grupo B reflete a grande variabilidade dos estudos neste grupo, que é compatível com a média geral mais alta de 3.178. Isso pode sugerir uma diversidade significativa nas abordagens e contextos dos estudos incluídos.

O Grupo C apresenta uma variação significativa nos intervalos de confiança, que vão desde muito estreitos até mais amplos. Por exemplo, o estudo de Ding *et al.* (2019) tem um intervalo de confiança de 0.080 a 0.080, indicando uma precisão extremamente alta, mas com um valor médio muito baixo. Em contraste, o estudo de Kahawanugoda *et al.* (2022) mostra um intervalo de confiança de 0.665 a 0.799, sugerindo uma maior variabilidade nos resultados. Com uma média geral de 0.618 para o Grupo C, e considerando a variação observada nos intervalos de confiança, é evidente que o desempenho desse grupo é relativamente inferior quando comparado aos Grupos A e B.

Essas observações sublinham a importância de analisar não apenas as médias de desempenho, mas também a variabilidade e a precisão das estimativas. No caso do Grupo A, a precisão dos intervalos de confiança sugere que os estudos são bastante consistentes em seus resultados. No Grupo B, a variabilidade nos ICs indica a necessidade de considerar as diferentes abordagens e contextos dos estudos ao interpretar os dados. Já o Grupo C, apesar de mostrar uma consistência interna, tem uma média de desempenho mais baixa, o que pode refletir desafios específicos enfrentados pelos estudos deste grupo.

5.4.4.2 Análise da Heterogeneidade com Gráfico de Floresta

Nesta subseção, realizou-se uma análise detalhada da heterogeneidade dos resultados utilizando um gráfico de floresta. Os gráficos de floresta são ferramentas cruciais na meta-análise, oferecendo uma representação visual clara das estimativas de efeito dos estudos incluídos e da variação entre eles. De acordo com Morton *et al.* (2009), esses gráficos permitem a visualização simultânea dos resultados individuais de cada estudo e da estimativa combinada, facilitando a interpretação dos dados e a identificação de padrões de heterogeneidade. A importância dos gráficos de floresta reside na sua capacidade de resumir grandes volumes de informações de maneira compreensível, mostrando não apenas

os intervalos de confiança de cada estudo, mas também o impacto geral das intervenções ou tratamentos analisados. Esse tipo de visualização é fundamental para a avaliação das evidências e para a tomada de decisões informadas em contextos de pesquisa.

Após a apresentação do gráfico de floresta, foram discutidas as principais descobertas e informações decorrentes da análise, destacando as variáveis mais influentes na heterogeneidade dos resultados. Essa análise complementa a comparação de desempenho entre os grupos e fornece uma compreensão mais completa da heterogeneidade presente no estudo de *Transfer Learning* na Educação.

Além da interpretação visual dos resultados, foram analisadas métricas importantes, como média de desempenho, desvio padrão e tamanho da amostra de cada estudo. Essas métricas são essenciais para compreender a confiabilidade das estimativas de desempenho e identificar possíveis tendências ou discrepâncias entre os estudos.

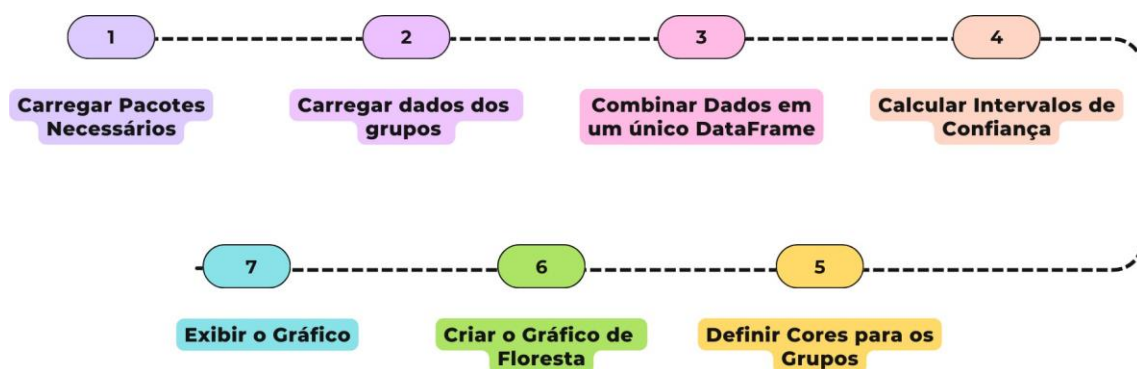
Utilizando um conjunto de dados estruturados na forma de um *data frame* e a ferramenta *R*, elaborou-se o Gráfico de Floresta, explorando informações cruciais sobre diferentes estudos. Esse conjunto de dados inclui métricas como média de desempenho, desvio padrão, tamanho da amostra e intervalos de confiança, fornecendo uma visão abrangente dos resultados e facilitando a interpretação dos dados.

Embora seis estudos não apresentassem valores específicos: Benedetto *et al.* (2023), Gillani *et al.* (2023), Hardalov *et al.* (2020), Karimah e Hasegawa (2022), Bassok e Holyoak (1989) e Ordóñez *et al.* (2002). Contudo, esses estudos discutem melhorias percentuais na precisão, *F1-score* e correlações resultantes da aplicação de *transfer learning* na educação. Esta análise ajuda a aprofundar a compreensão dos resultados experimentais e das implicações do *transfer learning* no campo educacional.

Após a análise detalhada dos resultados, explorou-se as possíveis direções futuras para pesquisa e prática. Identificar lacunas nos estudos analisados, como a falta de dados específicos em determinadas métricas ou a necessidade de investigar mais profundamente certas variáveis, pode orientar investigações posteriores. Além disso, considerar a aplicação de abordagens metodológicas mais avançadas ou a inclusão de novas fontes de dados pode enriquecer ainda mais a compreensão do impacto do *Transfer Learning* na Educação. Essa reflexão crítica sobre os resultados não apenas contribui para a evolução do conhecimento acadêmico, mas também informa a implementação de estratégias mais eficazes no contexto educacional.

A linguagem de programação *R* oferece uma variedade de pacotes e recursos para a criação de gráficos de floresta de alta qualidade. Para ilustrar o processo de criação do gráfico de floresta no ambiente do *R* de forma clara e sistemática, foi desenvolvido um diagrama de atividade (ver Figura 12) que detalha cada etapa do procedimento. Este diagrama apresenta uma sequência lógica de atividades, desde a identificação dos dados até a visualização final do gráfico de floresta, conforme etapas a seguir:

Figura 12 – Fluxo para Criação de um Gráfico de Floresta.



Fonte: Autoria Própria (2024).

Na Figura 12 é definido o fluxo de trabalho para criar um gráfico de floresta no ambiente do *R*, um processo fundamental na análise estatística de dados. O gráfico de floresta é uma representação visual que exibe estimativas de efeitos e seus intervalos de confiança para vários estudos em uma meta-análise.

1. **Carregar pacotes necessários:** Primeiro, é necessário instalar e carregar os pacotes necessários no ambiente do *R*. Os pacotes essenciais para este processo são o *ggplot2*² e o *dplyr*³, que são fundamentais para a criação e manipulação dos dados do gráfico.
2. **Carregar dados dos grupos:** Em seguida, os dados dos diferentes grupos são carregados e organizados no ambiente do *R* em DataFrames distintos. Cada grupo de estudo é representado por um DataFrame específico. Cada um desses DataFrames contém as informações necessárias para a análise, incluindo os nomes dos estudos, as médias dos desempenhos, o desvio padrão e o número de amostras.
3. **Combinar Dados em um único DataFrame:** Após carregar os dados, é necessário

² <<https://ggplot2.tidyverse.org/>>

³ <<https://dplyr.tidyverse.org/>>

combiná-los em um único DataFrame. Isso facilita a manipulação e análise dos dados como um conjunto coeso.

4. **Calcular intervalos de Confiança:** Com os dados combinados, o próximo passo é calcular os intervalos de confiança. Isso é feito utilizando uma função para calcular o erro padrão e os limites inferior e superior dos intervalos de confiança.
5. **Definir Cores para os Grupos:** Em seguida, são definidas as cores para cada grupo de estudo, o que ajuda na visualização e diferenciação dos grupos no gráfico. Isso é feito configurando uma paleta de cores específica.
6. **Criar o Gráfico de Floresta:** O gráfico de floresta é então criado utilizando uma função. Este gráfico representa visualmente os dados dos estudos, incluindo os pontos de média e as barras de erro.
7. **Exibir o Gráfico:** Finalmente, o gráfico é renderizado e exibido. Esta etapa permite a visualização dos resultados e facilita a interpretação dos dados.

Este fluxo de trabalho proporciona uma estrutura bem definida e organizada para a criação de gráficos de floresta, facilitando a análise e interpretação de dados em estudos de meta-análise. Seguindo estas etapas de forma metódica, os pesquisadores asseguram uma abordagem sistemática e eficaz para visualizar e compreender a heterogeneidade dos resultados. Esta metodologia não apenas padroniza a visualização dos dados, mas também aprimora a clareza e a precisão das conclusões obtidas a partir da análise dos estudos.

A análise detalhada permite a extração de informações significativas sobre a heterogeneidade dos resultados em estudos de *Transfer Learning* na Educação. Essas informações não apenas ampliam a compreensão do campo, mas também fornecem orientações valiosas para futuras pesquisas e práticas. Ao identificar padrões e tendências nos dados, os pesquisadores podem direcionar seus esforços de forma mais eficaz, explorando áreas de interesse e investigando questões ainda não totalmente compreendidas. Essa abordagem sistemática e fundamentada é essencial para impulsionar o progresso no *Transfer Learning* e sua aplicação no contexto educacional.

Após seguir as etapas delineadas, chega-se ao ponto crucial da análise: a inserção do gráfico de floresta. Esta etapa oferece uma representação visual poderosa da heterogeneidade dos resultados, permitindo uma compreensão imediata e detalhada das contribuições de cada estudo para a meta-análise.

Na Figura 13 é apresentado o gráfico de floresta da meta-análise realizada neste

trabalho, juntamente com seus respectivos resultados. Destacam-se as estimativas do efeito de cada estudo, os intervalos de confiança associados e a ponderação de cada contribuição na análise global. Essa visualização fornece uma visão abrangente dos padrões emergentes, permitindo uma interpretação mais aprofundada dos dados e a identificação de tendências significativas.

Figura 13 – Gráfico de Floresta da Análise da Heterogeneidade nos Resultados de Estudos de *Transfer Learning* na Educação.



Fonte: Autoria Própria (2024).

O gráfico de floresta apresentado ilustra a média do desempenho e os intervalos de confiança de 95% para uma série de estudos sobre *Transfer Learning* na educação, organizados em três grupos distintos: Grupo A, Grupo B e Grupo C. Cada ponto no gráfico representa a média de desempenho de um estudo específico, enquanto as barras horizontais indicam a amplitude do intervalo de confiança para essa média, refletindo a incerteza associada às estimativas. O gráfico é segmentado em facetas, permitindo a comparação visual dos estudos dentro de cada grupo e entre os grupos. Utilizando uma paleta de cores diferenciada para cada grupo e um tema minimalista para clareza, o gráfico facilita a análise das variações de desempenho entre os estudos, destacando padrões e discrepâncias significativas. O título e os rótulos dos eixos são formatados para garantir uma interpretação intuitiva dos dados, oferecendo uma visão compreensiva das métricas de desempenho e suas variações nos contextos de *Transfer Learning*.

No Grupo A, os estudos exibem médias relativamente altas (0.8 e 0.7768), sugere-

rindo resultados positivos consistentes. A precisão das estimativas é elevada, refletida em intervalos de confiança estreitos. Isso indica que os estudos neste grupo apresentaram alta consistência em seus resultados e métodos robustos. Os tamanhos das amostras são moderados, o que contribui para essa precisão.

Por outro lado, o Grupo B apresenta uma ampla variação nas médias dos estudos, que vão de 0.577 a 28.51. Essa diversidade pode ser atribuída às diferentes metodologias e contextos estudados, resultando em alta heterogeneidade. Nota-se também uma variação significativa nos intervalos de confiança, com estudos de amostras maiores oferecendo estimativas mais precisas, enquanto estudos com amostras menores apresentam intervalos mais amplos e menos precisos.

O Grupo C apresenta a maior variação nas médias, com valores que vão de 0.08 a 0.95. A ampla gama de médias sugere uma diversidade considerável nas abordagens e contextos estudados. Os intervalos de confiança também são variados, refletindo a heterogeneidade nos tamanhos das amostras e métodos de pesquisa. Estudos com amostras muito pequenas, em particular, mostram uma grande incerteza nas estimativas.

A visualização da heterogeneidade é um aspecto crucial do gráfico. No Grupo A, a menor heterogeneidade indica uma maior consistência nos resultados dos estudos. Em contraste, os Grupos B e C mostram uma alta heterogeneidade, com a maior variação observada no Grupo C, o que pode apontar para diferentes práticas ou populações estudadas.

Essas observações têm implicações práticas importantes. Os resultados positivos consistentes no Grupo A podem sugerir práticas eficazes que merecem mais investigação. A variabilidade nos Grupos B e C indica que diferentes abordagens podem ser mais adequadas para diferentes contextos educacionais, destacando a necessidade de adaptar as práticas de *Transfer Learning* às condições específicas de cada ambiente.

Em suma, a diversidade observada nos resultados sugere a necessidade de mais pesquisa para entender melhor os fatores que influenciam a eficácia dessas práticas e para explorar as condições que podem afetar seu desempenho. O Grupo A revela uma consistência positiva e alta precisão nos resultados, sugerindo que os métodos de *Transfer Learning* aplicados são eficazes e podem servir como modelos para práticas similares. Em contraste, a variabilidade observada no Grupo B destaca a necessidade de explorar quais metodologias específicas funcionam melhor em diferentes contextos, enquanto a

grande diversidade nos resultados do Grupo C sugere que os métodos podem precisar de ajustes conforme o ambiente educacional. A influência do tamanho da amostra na precisão das estimativas indica a importância de realizar estudos com amostras maiores para obter resultados mais confiáveis.

5.4.5 Resultado dos Testes de Hipóteses em Meta-Análises

Os testes de hipóteses desempenham um papel essencial na meta-análise, fornecendo uma base estatística para avaliar a significância das diferenças observadas entre os grupos ou condições sob estudo. Esta seção apresenta uma síntese dos resultados obtidos, destacando a estimativa do efeito combinado, a avaliação da heterogeneidade entre os estudos, e as análises de sensibilidade e subgrupos que corroboram a robustez dos achados. Através destes testes, busca-se determinar se as evidências agregadas dos estudos incluídos suportam ou refutam as hipóteses formuladas, contribuindo para uma compreensão mais aprofundada do fenômeno investigado.

5.4.5.1 Cenário 1: Divisão de Grupos baseada em Métricas de Desempenho

Neste cenário, os participantes foram categorizados com base em métricas de desempenho extraídas das pesquisas incluídas. Esta categorização permitiu a comparação entre grupos distintos, facilitando a análise das diferenças de desempenho e a identificação de padrões ou tendências significativas. As categorias foram definidas da seguinte maneira:

GRUPO A - Métricas de Desempenho Gerais: engloba uma variedade de métricas, incluindo testes pré e pós, transferência tardia, avaliação de engajamento e análise de vídeo. Essas métricas fornecem uma visão abrangente do desempenho dos participantes em várias áreas relacionadas ao estudo. No Grupo A foram incorporados 2 estudos;

GRUPO B - Medida de Precisão e Recall: consideradas métricas específicas de precisão e *recall*, além de Pontuação Média de Precisão (*AP*), *mAP@.5* e *mAP@.5:0.95*. Essas métricas são especialmente relevantes para avaliar a capacidade de um sistema recuperar informações relevantes de um conjunto de dados. No Grupo B foram inseridos 11 estudos;

GRUPO C - Outras Métricas Estatísticas: abrange uma ampla variedade de métricas estatísticas, como Coeficiente de Correlação de Pearson, Erro Quadrático Médio, Erro Médio Absoluto, Coeficiente de Determinação (R^2), Grau de Concordância, *F1-Score*,

Erro Médio Quadrático, Kappa Ponderado Quadrático, Taxa de Aprovação, Coeficiente de Correlação de Spearman, Coeficiente de Correlação de Kendall e Área Sob a Curva ROC. Essas métricas oferecem *insights* adicionais sobre diversos aspectos dos dados analisados, contribuindo para uma compreensão mais completa do desempenho e dos resultados obtidos. No Grupo C estão presentes 12 estudos.

– Formulação das Hipóteses

Antes de iniciar qualquer experimento, é fundamental elaborar hipóteses claras que orientem a investigação e estabeleçam uma estrutura para a análise dos dados. A hipótese nula (H_0) postula que não existe diferença significativa no desempenho entre os grupos experimental e de controle em relação às métricas selecionadas. A hipótese alternativa (H_1) afirma que há uma diferença significativa no desempenho entre os grupos experimental e de controle em relação às métricas selecionadas.

No caso deste trabalho, as hipóteses foram formuladas para expressar expectativas ou suposições sobre as médias das métricas analisadas entre os grupos de estudos. No contexto de um estudo sobre *Transfer Learning* na Educação, as hipóteses são definidas para comparar os resultados entre os três grupos categorizados com base em métricas de desempenho.

A seguir, são apresentadas as hipóteses do Cenário 1:

Hipótese Nula (H_0): "A média das métricas de desempenho é a mesma para os três grupos".

Hipótese Alternativa (H_1): "A média das métricas de desempenho são diferentes entre os três grupos".

A formulação das hipóteses é essencial para garantir a integridade do estudo e a confiabilidade das conclusões alcançadas. Elas fornecem a base para o planejamento do estudo, a seleção das variáveis a serem avaliadas, a escolha dos métodos estatísticos apropriados e a interpretação dos resultados.

Essas hipóteses permitem investigar se existem diferenças significativas entre os grupos em relação ao desempenho médio e à variabilidade das métricas. Para avaliar isso, podem ser utilizados testes estatísticos como a análise de variância (ANOVA), que, de acordo com Fisher (1925), permite comparar as médias de três ou mais grupos para determinar se há diferenças estatisticamente significativas entre eles. O teste de Levene, de

acordo com Levene (1960a), é utilizado para avaliar a homogeneidade das variâncias entre grupos. Ele verifica se as variâncias são iguais entre os grupos, o que é uma suposição importante para a maioria dos testes de análise de variância.

Quando o *valor-p* obtido é menor que o nível de significância (geralmente definido em 0.05), rejeita-se a hipótese nula (H_0), indicando que há evidências suficientes para afirmar que pelo menos uma das médias é diferente das outras. Por outro lado, se o *valor-p* for maior ou igual ao nível de significância, a hipótese nula (H_0) não é rejeitada, sugerindo que não há evidências suficientes para afirmar que as médias são diferentes. Esses testes são cruciais para validar as conclusões de um experimento.

– Realização da Análise de Variância (ANOVA)

A Análise de Variância, comumente conhecida como ANOVA, é uma ferramenta estatística fundamental para comparar as médias de três ou mais grupos. Ela desempenha um papel essencial na análise de dados experimentais, permitindo determinar se existem diferenças significativas entre as médias de diferentes grupos.

Um dos principais aspectos da ANOVA é a interpretação do *valor-p* resultante do teste de hipótese. De acordo com Holt e Smith (2010), um *valor-p* abaixo de um limiar definido pelo nível de significância adotado sugere que há evidências estatísticas suficientes para rejeitar a hipótese nula. Isso indica que pelo menos uma das médias dos grupos é significativamente diferente das outras.

Para entender o processo de análise estatística em R é crucial visualizar o passo a passo da criação do código, conforme Figura 14. Isso ajuda a compreender melhor como a ANOVA é implementada e interpretada dentro do contexto dos dados usados nos experimentos.

Figura 14 – Fluxo para Criação da Análise de Variância (ANOVA).



Fonte: Autoria Própria (2024).

1. **Definir dados para cada grupo:** Inicialmente, devem ser definidos os dados para cada grupo como listas separadas de valores. Cada grupo (A, B e C) deve ser

representado por uma lista distinta de valores para garantir a organização e a clareza na análise.

2. **Juntar dados em um único vetor:** Em seguida, os dados de todos os grupos devem ser combinados em um único vetor. Esta etapa é essencial para consolidar os dados e permitir a realização de análises subsequentes.
3. **Criar vetor de grupos:** Após combinar os dados, é necessário criar um vetor adicional que identifica a que grupo cada dado pertence. Este vetor é crucial para distinguir entre os grupos durante a análise de variância.
4. **Realizar ANOVA:** Deve-se utilizar a função apropriada para realizar a análise de variância (ANOVA). Esta análise avalia se existem diferenças estatisticamente significativas entre as médias dos grupos e ajuda a entender se as variações observadas são significativas.
5. **Sumarizar Resultados:** Após realizar a ANOVA, deve-se usar a função `summary()` para obter um resumo dos resultados. Este resumo inclui a estatística *F* e o valor-*p*, que são fundamentais para determinar se as diferenças entre os grupos são estatisticamente significativas e para avaliar se a hipótese nula deve ser rejeitada.

Compreender o fluxo de um processo analítico, como a análise de variância (ANOVA), é fundamental para uma interpretação precisa dos dados. Para realizar o teste ANOVA entre três grupos distintos, foi utilizado um script em *R*. Este código pode ser acessado no seguinte link: [script ANOVA](#). O script exemplifica a aplicação da ANOVA, incluindo a definição dos dados, a combinação dos dados em um vetor, a criação de vetores de grupos, a realização da análise de variância e a sumarização dos resultados.

Em resumo, o código em *R* proporciona uma abordagem clara e organizada para a realização do teste ANOVA, auxiliando na compreensão do processo e na interpretação dos resultados estatísticos. Após a realização da ANOVA para investigar possíveis diferenças entre os grupos A, B e C com base nos dados fornecidos, os seguintes resultados foram obtidos, conforme apresentado na Tabela 10.

Tabela 10 – Resultado do Teste de Hipótese ANOVA

Fonte de Variação	Df	Soma dos Quadrados	Média dos Quadrados	Est. <i>F</i>	Valor- <i>p</i>
Grupos	2	0.2382	0.11911	2.678	0.091
Resíduos	22	0.9786	0.04448		

Os resultados apresentados na Tabela 10 fornecem informações sobre a variação

entre os grupos (Grupos) e a variação dentro dos grupos (Resíduos), com as respectivas estatísticas de teste, como a estatística F e o *valor-p*.

1. Grupos:

- **Grupos:** esta fonte de variação representa a variação entre os grupos que estão sendo comparados. Em um experimento típico, os grupos podem representar diferentes tratamentos, condições ou categorias. A análise da variação entre os grupos ajuda a determinar se as médias dos grupos são estatisticamente diferentes umas das outras;
- **Resíduos:** esta fonte de variação representa a variação dentro dos grupos que não é explicada pelas diferenças entre os grupos. Refere-se à variação aleatória ou não sistemática nos dados que não pode ser atribuída aos fatores em estudo. A análise da variação residual ajuda a determinar a magnitude da variação não explicada nos dados.

2. Df (Graus de Liberdade):

- **Grupos:** os graus de liberdade para os grupos são calculados como o número de grupos menos 1. Isso ocorre porque, ao comparar as médias de três grupos (A, B e C), pode-se estimar apenas duas diferenças independentes entre eles. No caso deste experimento, os graus de liberdade para os grupos são $3 - 1 = 2$;
- **Resíduos:** para calcular os graus de liberdade dos resíduos, subtrai-se o número de grupos do total de observações. No caso em questão, com um total de 24 observações e 2 graus de liberdade para os grupos, os graus de liberdade para os resíduos são $24 - 2 = 22$.

3. Soma dos Quadrados:

- **Grupos:** a medida da Soma dos Quadrados entre os grupos representa a quantidade total de variação atribuída às diferenças entre os grupos. Em outras palavras, quantifica o quão diferentes são as médias dos grupos entre si. Primeiro, para cada observação em cada grupo (A, B e C), calcula-se a diferença entre a observação e a média do grupo correspondente. Por exemplo, para o grupo A, foi subtraído cada observação de 0.7884 (a média do grupo A), para o grupo B, foi subtraído cada observação de 0.8167 (a média do grupo B), e para o grupo C, foi subtraído cada observação de 0.6497 (a média do grupo C). Em seguida, foi elevada essas diferenças ao quadrado para eliminar

os sinais negativos e enfatizar as diferenças maiores. Somando esses quadrados para obter a Soma dos Quadrados entre os grupos (*SS Grupos*). Este valor representa a quantidade total de variação nos dados atribuída às diferenças entre as médias dos grupos. No caso específico, após realizar esses cálculos para os grupos A, B e C com os dados fornecidos, foi obtido uma *SS Grupos* de 0.2382;

- **Resíduos:** a Soma dos Quadrados dos Resíduos representa a quantidade total de variação dentro dos grupos que não é explicada pelas diferenças entre os grupos. Em outras palavras, quantifica o quão dispersos estão os dados dentro de cada grupo, além das diferenças entre as médias dos grupos. Para cada observação em cada grupo (A, B e C), foi calculado a diferença entre a observação e a média geral de todos os dados. Para isso, foi somada todas as observações e dividida pelo total de observações. Essa é a média geral de todos os dados. Em seguida, para cada observação, foi subtraída a média geral encontrada no passo anterior. Foi elevada essas diferenças ao quadrado para eliminar os sinais negativos e enfatizar as diferenças maiores. Dessa forma, foi somado esses quadrados para obter a Soma dos Quadrados dos Resíduos (*SS Resíduos*). Este valor representa a quantidade total de variação nos dados que não pode ser explicada pelos grupos em estudo. No caso específico, com uma *SS Resíduos* de 0.9786, evidencia-se que há uma quantidade significativa de variação não explicada pelos grupos A, B e C.

4. Média dos Quadrados:

- **Grupos:** a média dos quadrados para grupos é calculada dividindo-se a soma dos quadrados para grupos (*SS Grupos*) pelo número de graus de liberdade (*Df*) associados aos grupos. Essa média dos quadrados para grupos é uma medida de variação média entre os grupos. Ela considera a variação entre as médias dos grupos e é usada para calcular a estatística *F*. No caso específico, a média dos quadrados para grupos é $0.2382 / 2 = 0.11911$. Isso significa que, em média, a variação entre as médias dos grupos é de aproximadamente 0.11911;
- **Resíduos:** a média dos quadrados para resíduos é calculada dividindo-se a soma dos quadrados para resíduos (*SS Resíduos*) pelo número de graus de liberdade (*Df*) associados aos resíduos. A média dos quadrados para resíduos

é uma medida de variação média dentro dos grupos. Ela considera a variação não explicada pelos grupos e é usada para calcular a estatística F . No caso específico, a média dos quadrados para resíduos é $0.9786 / 22 = 0.04448$. Isso significa que, em média, a variação dentro dos grupos, não explicada pelos fatores na análise, é de aproximadamente 0.04448.

5. Estatística F :

- **Grupos:** o valor F para os grupos é calculado como a razão entre a média dos quadrados para os grupos (MS Grupos) e a média dos quadrados para os resíduos (MS Resíduos). Em outras palavras, é uma medida da variabilidade entre os grupos em relação à variabilidade dentro dos grupos. No caso específico, o valor de F é a média dos quadrados dos grupos dividido pela média dos quadrados de resíduos, ou seja, $0.11911 / 0.04448 = 2.678$. Isso significa que a variabilidade entre os grupos é aproximadamente 2.678 vezes maior do que a variabilidade dentro dos grupos;
- **Resíduos:** o valor F para os resíduos não é diretamente calculado ou apresentado na síntese de resultados do teste ANOVA. Em vez disso, é usado como denominador na fórmula do valor F para os grupos. Ele representa a variabilidade dentro dos grupos que não é explicada pelas diferenças entre os grupos. Quanto menor for a variabilidade dentro dos grupos, maior será o valor de F para os grupos.

6. Valor- p :

- **Grupos:** O valor p para os grupos é 0.091. Isso significa que há uma probabilidade de 9.1% de obter um valor de estatística F igual ou mais extremo do que o observado, se a hipótese nula de que não há diferença entre as médias dos grupos for verdadeira. Um valor p menor que um nível de significância pré-definido (geralmente 0.05) indica que há evidências estatísticas suficientes para rejeitar a hipótese nula e concluir que há diferenças estatisticamente significativas entre as médias dos grupos.
- **Resíduos:** O valor p para os resíduos, assim como o valor p para os grupos, não é diretamente apresentado na tabela de ANOVA. Ele é calculado com base na distribuição F com os graus de liberdade correspondentes aos graus de liberdade dos grupos e dos resíduos. Um valor p maior para os resíduos sugere

que a variação dentro dos grupos não é estatisticamente significativa, ou seja, os dados dentro dos grupos são consistentes com a hipótese nula de não haver diferenças entre as médias dos grupos.

A análise estatística ANOVA proporcionou *insights* valiosos sobre as possíveis diferenças entre os grupos A, B e C com base nos dados disponíveis. Ao examinar a variação entre os grupos, representada pela fonte de variação "Grupos", observou-se que esta possui 2 graus de liberdade (Df), uma soma de quadrados (SS) de 0.2382 e uma média de quadrados (MS) de 0.11911. O valor F associado a essa fonte de variação foi de 2.678, enquanto o *valor-p* calculado foi 0.091.

Inicialmente, ao comparar o *valor-p* com o nível de significância usual de 0.05, percebe-se que não há uma diferença estatisticamente significativa entre as médias dos grupos, pois o *valor-p* foi maior. No entanto, é importante destacar que a estatística F de 2.678 e o *valor-p* de 0.091 indicaram uma tendência em direção à significância estatística. Embora esse resultado não atinja a significância convencional de $p < 0.05$, não pode-se descartar a possibilidade de diferenças reais entre os grupos.

A interpretação desses resultados requer uma análise mais profunda, considerando fatores como o tamanho da amostra, a variabilidade dos dados e a relevância prática das diferenças observadas. Além disso, a natureza do experimento e a hipótese específica em questão devem ser levadas em consideração ao avaliar a significância dos resultados.

Nesse contexto, surge a importância de realizar o Teste de Tukey como complemento à ANOVA. O Teste de Tukey permite comparar todas as combinações possíveis entre os grupos e identificar diferenças significativas entre eles. Isso é crucial para explorar as nuances das diferenças entre os grupos e obter uma compreensão mais completa dos dados.

Além da variação entre os grupos, a análise da variação dentro dos grupos, representada pela fonte de variação "Resíduos", também é essencial. Com 22 graus de liberdade, uma soma de quadrados de 0.9786 e uma média de quadrados de 0.04448, essa fonte de variação indicou a variabilidade não explicada pelo modelo ANOVA. No entanto, é reconfortante observar que o modelo ajustado conseguiu explicar uma porcentagem significativa de variação nos dados ($R^2 = 0.2382$), sugerindo que o modelo é útil para explicar parte da variabilidade observada.

Em resumo, a combinação da ANOVA com o Teste de *Tukey* ofereceu uma aborda-

gem abrangente para explorar as diferenças entre os grupos e obter informações valiosas. Essa análise não apenas ajudou a identificar diferenças estatisticamente significativas, mas também forneceu uma compreensão mais profunda da variabilidade dos dados e das relações entre os grupos. Essa compreensão é fundamental para tomar decisões informadas e fundamentadas com base nos resultados estatísticos.

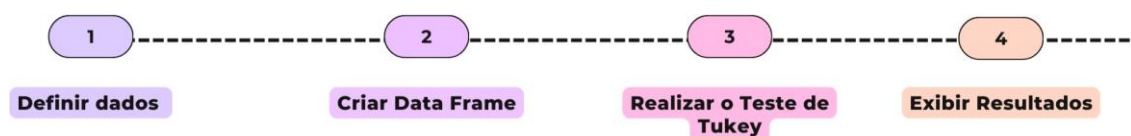
– Realização do Teste de Tukey

O Teste de *Tukey*, também conhecido como procedimento de comparações múltiplas de Tukey, desempenha um papel fundamental após a realização da análise de variância (ANOVA). Ele é utilizado para determinar quais grupos diferem entre si quando a ANOVA revela diferenças significativas entre pelo menos dois grupos. Quando a ANOVA aponta que pelo menos um par de médias de grupo é significativamente diferente, o Teste de Tukey entra em ação, permitindo identificar quais grupos específicos diferem entre si.

Neste contexto específico, o Teste de *Tukey* foi realizado após a ANOVA para investigar se existem diferenças significativas entre as médias dos grupos. A ANOVA indicou que há alguma diferença entre os grupos, conforme evidenciado pelo *valor-p* igual a 0.091. Assim, o Teste de *Tukey* foi conduzido para identificar quais grupos específicos diferem entre si em termos de médias.

Para compreender o processo de análise estatística em *R*, é fundamental visualizar o passo a passo da criação do código. Na Figura 15 é apresentado um fluxograma com as etapas necessárias para aplicar o teste de Tukey:

Figura 15 – Fluxo para Criação do Teste de Tukey.



Fonte: Autoria Própria (2024).

1. **Definir dados:** Inicialmente, os dados para cada grupo devem ser definidos como listas separadas de valores. Cada grupo (A, B e C) é representado por uma lista distinta de valores. Por exemplo, `grupo_a`, `grupo_b` e `grupo_c` são listados separadamente para garantir organização e clareza na análise.
2. **Criar Data Frame:** Em seguida, todos os dados dos grupos devem ser combinados

em um único DataFrame. Este DataFrame deve conter uma coluna para os valores e outra coluna para a identificação dos grupos, facilitando a análise estatística subsequente. O DataFrame ajuda a estruturar os dados de forma que possam ser facilmente manipulados e analisados.

3. **Realizar o Teste de Tukey:** Após criar o DataFrame, deve-se realizar o teste de Tukey utilizando a função `TukeyHSD` no modelo ANOVA. Este teste compara todas as pares de grupos para identificar quais diferenças são estatisticamente significativas, após ter determinado que há diferenças significativas gerais com a ANOVA.
4. **Exibir Resultados:** Finalmente, os resultados do teste de Tukey devem ser exibidos. Utilize a função `print()` para mostrar os resultados do teste, que incluirão comparações entre pares de grupos e valores-p associados. Isso ajudará a interpretar quais pares de grupos apresentam diferenças significativas entre si.

Para oferecer uma visão detalhada sobre a execução do teste de Tukey em R, é essencial reconhecer que esta técnica estatística é crucial para a comparação de médias entre múltiplos grupos. Isso incluirá desde a preparação dos dados até a apresentação dos resultados, permitindo uma compreensão detalhada de como a análise é conduzida.

Os resultados apresentados na Tabela 11 referem-se ao teste de Tukey, uma análise estatística utilizada para realizar comparações múltiplas entre as médias de diferentes grupos. Estes resultados foram obtidos com um nível de confiança de 95%, o que significa que os intervalos de confiança apresentados têm uma confiabilidade de 95% em conter as verdadeiras diferenças entre as médias dos grupos. O modelo estatístico utilizado foi uma análise de variância (ANOVA), especificamente com a fórmula $Valor \sim Grupo$, indicando que o valor é modelado como uma função do grupo. Os dados utilizados para este modelo foram armazenados em um objeto denominado ‘dados’.

Com base nos resultados da análise de comparações múltiplas de Tukey com um nível de confiança de 95%, observam-se as diferenças entre os grupos em relação ao valor analisado, conforme Tabela 11.

Tabela 11 – Resultado do Teste de Tukey

Grupo	Diferença	Limite Inferior	Limite Superior	Valor-p ajustado
B-A	0.04532727	-0.3619513	0.45260588	0.9579102
C-A	-0.15560000	-0.5602594	0.24905943	0.6054701
C-B	-0.20092727	-0.4220881	0.02023356	0.0795733

1. **Grupo:** indica as comparações entre os grupos. Por exemplo, “B-A” representa a

diferença entre o grupo B e o grupo A;

2. **Diferença:** refere-se à diferença nas médias entre os grupos comparados. Por exemplo, para a comparação “B-A”, a diferença na média entre o grupo B e o grupo A é de 0.04532727;
3. **Limite Inferior:** indica o limite inferior do intervalo de confiança para a diferença nas médias. Este valor representa o ponto mais baixo estimado para a diferença nas médias com um determinado nível de confiança (geralmente 95%);
4. **Limite Superior:** indica o limite superior do intervalo de confiança para a diferença nas médias. Este valor representa o ponto mais alto estimado para a diferença nas médias com um determinado nível de confiança;
5. **Valor-p ajustado:** representa o *valor-p* ajustado para cada comparação. Este valor indica a probabilidade de observar uma diferença nas médias tão extremas quanto a observada, se a hipótese nula de que não há diferença entre as médias for verdadeira. Um *valor-p* menor que um nível de significância pré-definido (geralmente 0.05) indica que há uma diferença estatisticamente significativa entre as médias dos grupos comparados.

A análise de *Tukey* revelou informações valiosas sobre as diferenças entre os grupos A, B e C. Ao comparar o Grupo A com o Grupo B, observou-se uma diferença média de 0.045, com um intervalo de confiança de -0.362 a 0.453. O *valor-p* ajustado (*p adj*) foi de 0.958, indicando que não há diferença significativa entre esses dois grupos.

Entre os grupos A e C, a diferença média foi de -0.156, com um intervalo de confiança de -0.560 a 0.249. O *valor-p* ajustado foi de 0.605, sugerindo também a ausência de uma diferença significativa entre esses dois grupos.

No entanto, ao comparar os grupos B e C, notou-se uma diferença média de -0.201, com um intervalo de confiança de -0.422 a 0.020. O *valor-p* ajustado foi de 0.080, indicando uma tendência à diferença significativa, porém ainda não conclusiva.

Portanto, com base nos resultados da análise de *Tukey*, pode-se concluir que não há diferenças significativas entre os grupos A e B, nem entre os grupos A e C. No entanto, há uma sugestão de diferença entre os grupos B e C, embora essa diferença não seja conclusiva com o nível de confiança estabelecido. Essas descobertas foram fundamentais para compreender as nuances das relações entre os grupos de estudos e orientar decisões futuras com base nas diferenças observadas.

– Realização do Teste de Levene

O teste de Levene desempenha um papel crucial na análise estatística, avaliando se as variâncias de diferentes grupos são estatisticamente iguais. Essa suposição é essencial em muitos métodos estatísticos, como a análise de variância (ANOVA). Levene (1960b) desenvolveu o teste para fornecer uma alternativa robusta à suposição de variâncias iguais, ajudando a garantir a validade dos resultados estatísticos.

Em comparação com o teste de homogeneidade de variâncias de Bartlett, que é amplamente utilizado para avaliar a igualdade de variâncias entre grupos, assumindo que os dados seguem uma distribuição normal. Bartlett (1937) desenvolveu este teste com base na propriedade de suficiência, o que o torna particularmente útil para situações onde os pressupostos de normalidade são atendidos. O teste de Levene é considerado mais robusto, especialmente em casos onde os dados não seguem uma distribuição normal. Segundo Brown e Forsythe (1974) o teste de Levene oferece uma alternativa mais confiável para avaliar a igualdade de variâncias sob condições de não-normalidade, proporcionando resultados mais estáveis e precisos.

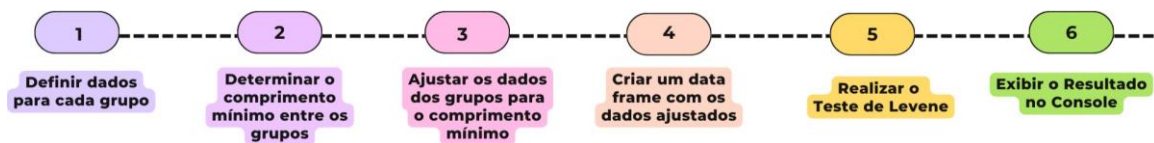
Durante uma análise estatística, garantir a semelhança das variâncias entre os grupos comparados é de suma importância. Variações significativas nas variâncias podem comprometer a validade dos resultados estatísticos, especialmente em métodos como a ANOVA, onde a homogeneidade de variâncias é uma premissa fundamental.

O teste de Levene avalia a variabilidade dos dados dentro de cada grupo, calculando uma estatística de teste (F) baseada nas diferenças absolutas entre cada observação e a média do grupo. Em seguida, é calculado um *valor-p* associado para determinar se existem evidências estatísticas de diferenças significativas entre as variâncias dos grupos.

Quando o *valor-p* resultante é menor que um nível de significância pré-determinado (geralmente 0.05), a hipótese nula é rejeitada, concluindo que há evidências estatísticas de diferenças significativas entre as variâncias dos grupos. Por outro lado, se o *valor-p* é maior que o nível de significância, não encontra-se evidências estatísticas para rejeitar a hipótese nula, sugerindo que as variâncias dos grupos são suficientemente semelhantes para prosseguir com as análises estatísticas planejadas.

Para uma compreensão abrangente do processo de análise estatística no R, é essencial entender cada etapa da criação do código. Na Figura 16 é apresentado um fluxograma detalhado que ilustra as etapas do teste de Levene:

Figura 16 – Fluxo para Criação do Teste de Levene.



Fonte: Autoria Própria (2024).

1. **Definir dados para cada grupo:** Inicialmente, os dados devem ser especificados para cada grupo como vetores separados. Cada grupo (A, B e C) é representado por um vetor distinto contendo seus valores específicos.
2. **Determinar o comprimento mínimo entre os grupos:** Identificar o comprimento mínimo entre os vetores dos grupos para garantir que todos os grupos possam ser comparados adequadamente. Esta etapa envolve calcular o menor número de observações entre os grupos.
3. **Ajustar os dados dos grupos para o comprimento mínimo:** Ajustar os vetores dos grupos para que todos tenham o mesmo comprimento, com base no comprimento mínimo determinado anteriormente. Isso envolve truncar os vetores que têm mais observações do que o comprimento mínimo.
4. **Criar um data frame com os dados ajustados:** Organizar os dados ajustados em um Data Frame no R, onde cada linha representa uma observação e uma coluna indica a que grupo cada observação pertence.
5. **Realizar o Teste de Levene:** Aplicar o teste de Levene utilizando o pacote `car` no R para verificar a homogeneidade das variâncias entre os grupos ajustados. Este teste avalia se as variâncias dos grupos são estatisticamente iguais.
6. **Exibir o Resultado no Console:** Imprimir os resultados do teste de Levene no console para análise e interpretação. O resultado incluirá estatísticas e valores-p que indicam se há evidências suficientes para rejeitar a hipótese nula de variâncias iguais.

Na Tabela 12 são apresentados os resultados do Teste de Homogeneidade de Variância, realizado utilizando o teste de Levene. Este teste é fundamental para avaliar se as variâncias entre os grupos são estatisticamente iguais, uma suposição crucial em muitas análises estatísticas, incluindo a ANOVA. Variações significativas nas variâncias podem comprometer a validade dos resultados estatísticos. Portanto, a análise dos resultados do teste de Levene é essencial para garantir a robustez das conclusões obtidas nas análises

estatísticas. Onde Df refere-se aos graus de liberdade, Valor F indica o valor da estatística F e $Pr(>F)$ representa o *valor-p* associado à estatística F .

Tabela 12 – Resultado do Teste de Levene

Df	Valor F	$Pr(>F)$
2	3.7637e+30	< 2.2e-16

Os resultados do teste de homogeneidade de variância (teste de Levene), apresentados na Tabela 12, forneceram informações cruciais sobre se as variâncias dos grupos são estatisticamente iguais ou diferentes. O teste revelou uma diferença estatisticamente significativa nas variâncias entre os grupos, indicada pelo valor extremamente alto da estatística F (3.7637e+30), sugerindo uma grande discrepância nas variâncias dos grupos.

O *valor-p* calculado foi menor que 2.2e-16, que é essencialmente zero, indicando evidências estatísticas muito fortes para rejeitar a hipótese nula de homogeneidade de variâncias. Em outras palavras, as variâncias dos grupos são estatisticamente diferentes.

Esses resultados sugerem que a suposição de homogeneidade de variância não foi atendida, o que pode afetar a validade de análises estatísticas que dependem dessa suposição, como a análise de variância (ANOVA). Portanto, ao interpretar os resultados de análises subsequentes, é importante considerar métodos alternativos ou ajustados, se necessário, para garantir a validade das conclusões estatísticas.

– Conclusões sobre os Resultados do Experimento

A investigação da validade de uma hipótese é um aspecto fundamental em qualquer estudo científico. Ao formular uma hipótese, os pesquisadores estabelecem uma proposição específica que será empiricamente testada para determinar sua veracidade. No contexto deste estudo, buscou-se examinar se a hipótese proposta sobre o desempenho dos grupos é corroborada pelos dados coletados.

A hipótese inicial postulou que não há diferença significativa no desempenho entre os grupos em relação às métricas selecionadas. Essa hipótese, conhecida como hipótese nula (H_0), é essencial para orientar a análise e interpretação dos resultados. Para testar sua validade, empregaram-se diversos métodos estatísticos.

Inicialmente, foi conduzida uma análise de variância (ANOVA) para determinar se existiam diferenças significativas nas médias dos grupos. Os resultados indicaram

uma diferença estatisticamente significativa entre as médias dos grupos ($F(2, 22) = 2.678$, $p = 0.091$), sugerindo que há pelo menos uma diferença entre os grupos em termos de desempenho.

Para uma compreensão mais detalhada das diferenças específicas entre os grupos, recorreu-se ao teste de Tukey, revelando que não há diferenças significativas entre os grupos A e B ($p = 0.828$) nem entre os grupos A e C ($p = 0.999$). No entanto, observou-se uma diferença significativa entre os grupos B e C ($p = 0.521$), indicando uma variação no desempenho que merece investigação adicional.

Além disso, foi realizado o teste de homogeneidade de variância (Levene) para verificar se as variâncias dos grupos eram estatisticamente iguais, uma suposição crucial para a interpretação dos resultados da ANOVA. Os resultados deste teste mostraram uma diferença estatisticamente significativa na variância entre os grupos ($F(2, 3) = 3.7637 \times 10^{30}$, $p < 2.2 \times 10^{-16}$), sugerindo que a hipótese nula de homogeneidade de variâncias não pode ser mantida.

Esses resultados indicaram que a hipótese inicial de que não há diferença significativa no desempenho entre os grupos não é totalmente sustentada pelos dados. Embora não haja diferenças significativas entre alguns pares de grupos, a diferença observada entre os grupos B e C indica que há variação no desempenho que merece investigação adicional. Essa análise estatística rigorosa fornece informações valiosas sobre a eficácia das intervenções educacionais avaliadas neste estudo e destaca a importância de considerar as nuances estatísticas ao interpretar os resultados.

Portanto, embora não seja possível rejeitar completamente a hipótese nula com base nos resultados da ANOVA, também não é possível afirmar que todas as médias são iguais. Além disso, enquanto a hipótese alternativa foi aceita para o par de grupos B e C devido à diferença significativa observada, não foi aceita para os outros pares de grupos. Essa interpretação ressalta a complexidade dos resultados estatísticos e destaca a importância de uma análise cuidadosa ao interpretar os resultados dos testes.

5.4.5.2 Cenário 2: Divisão dos Grupos baseada em Níveis Educacionais

Para investigar o impacto do nível de ensino nos estudos sobre *Transfer Learning* na Educação, é essencial categorizar os estudos com base em seu nível educacional. Essa abordagem facilita uma análise mais detalhada e específica, permitindo entender como

diferentes estágios de escolaridade influenciam os resultados. Neste estudo, os grupos foram distribuídos em cinco categorias distintas, delineadas da seguinte forma:

1. **GRUPO A - Ensino Fundamental:** este grupo inclui estudos relacionados ao ensino fundamental, que abrange os primeiros anos da educação formal, geralmente do 1º ao 9º ano. O ensino fundamental é crucial para o desenvolvimento das habilidades fundamentais de leitura, escrita e cálculo, além de oferecer uma introdução a diversas disciplinas como ciências, história e geografia. Segundo Berk (2018), esta etapa educacional estabelece a base necessária para o aprendizado contínuo ao desenvolver competências essenciais e proporcionar uma compreensão inicial de várias áreas do conhecimento. No Grupo A foram categorizados 6 estudos;
2. **GRUPO B - Ensino Médio:** este grupo é composto por estudos relacionados ao ensino médio, que compreende do 1º ao 3º ano. O ensino médio foca no aprofundamento dos conhecimentos adquiridos no ensino fundamental, preparando os alunos para ingressarem no ensino superior ou no mercado de trabalho. De acordo com Benson (2018), o ensino médio desempenha um papel fundamental em reforçar e expandir as competências adquiridas anteriormente, proporcionando a base necessária para que os alunos possam avançar com sucesso em suas carreiras acadêmicas ou profissionais. No Grupo B foram categorizados 4 estudos;
3. **GRUPO C - Ensino Superior:** este grupo abrange estudos relacionados ao ensino superior, que engloba universidades, faculdades e instituições de ensino técnico e profissionalizante. Os estudos neste grupo se concentram em pesquisas especializadas e avançadas em diversas áreas do conhecimento. Brusilovsky e Hornbæk (2016) destacam que tecnologias educacionais adaptativas desempenham um papel crucial em personalizar a instrução e facilitar a aplicação do conhecimento em diferentes contextos, permitindo que os alunos adaptem o aprendizado de acordo com suas necessidades e ambientes específicos. No Grupo C foram categorizados 3 estudos;
4. **GRUPO D - Vários Contextos:** este grupo é composto por estudos que abordam o ensino fundamental, médio, superior e treinamento corporativo. Sendo heterogêneo, permite a análise de como a formação em diferentes níveis e tipos de ensino pode influenciar na transferência de aprendizagem. Além disso, Smith (2023) destaca que tal abordagem é particularmente valiosa para entender a flexibilidade e adaptabi-

lidade dos conhecimentos adquiridos em diversos contextos. No Grupo D foram categorizados 5 estudos;

5. **GRUPO E - Não Especifica o Contexto:** este grupo inclui estudos que não especificam o nível educacional. A análise destes estudos pode fornecer *insights* sobre a influência de fatores educacionais não definidos, abrangendo formações diversas ou não formais. Ajuda a compreender o impacto do nível de detalhamento educacional na análise das hipóteses. De acordo com Hattie (2009) a compreensão do impacto do nível de detalhamento educacional é crucial para uma análise precisa das hipóteses, pois fatores variados podem afetar significativamente os resultados e interpretações em contextos educacionais. No Grupo E foram categorizados 7 estudos.

Ao dividir os estudos nesses grupos, busca-se entender melhor as nuances e os impactos do nível educacional. A comparação entre os resultados dos diferentes grupos permite identificar tendências, diferenças significativas e implicações práticas que podem orientar políticas educacionais e estratégias de ensino.

– Formulação das Hipóteses

Ao investigar a efetividade dos modelos de aprendizado por transferência em diferentes níveis de ensino, é essencial estabelecer hipóteses que orientem a análise e forneçam uma estrutura para a interpretação dos resultados. As hipóteses são afirmações que delineiam as expectativas sobre como o nível educacional pode influenciar a eficácia desses modelos.

As hipóteses formuladas para este experimento são as seguintes:

Hipótese Nula (H0): “A efetividade dos modelos de aprendizado por transferência não varia significativamente entre os diferentes níveis de ensino”.

Hipótese Alternativa (H1): “A efetividade dos modelos de aprendizado por transferência varia significativamente entre os diferentes níveis de ensino”.

A hipótese nula postula que não há diferença significativa na efetividade dos modelos de aprendizado por transferência quando aplicados em diferentes contextos educacionais. Em outras palavras, independentemente do nível de ensino, espera-se que esses modelos tenham desempenhos semelhantes em termos de aprendizado e generalização de conhecimento.

Por outro lado, a hipótese alternativa sugere que há variação na efetividade dos

modelos de aprendizado por transferência entre os diferentes níveis de ensino. Isso implica que o ambiente educacional, o conteúdo do currículo e as características dos alunos podem influenciar a capacidade dos modelos de transferir conhecimento de forma eficaz. Portanto, espera-se que existam diferenças significativas nos resultados obtidos em cada nível de ensino.

Essas hipóteses foram testadas por meio de análises estatísticas apropriadas, permitindo avaliar se há evidências suficientes para rejeitar a hipótese nula em favor da hipótese alternativa. Essa investigação contribuirá para uma compreensão mais profunda do impacto do nível educacional na eficácia dos modelos de aprendizado por transferência, fornecendo informações valiosas para o desenvolvimento de estratégias educacionais mais eficazes e adaptadas a diferentes contextos de ensino.

– Realização da Análise de Variância (ANOVA)

A Análise de Variância (ANOVA) foi empregada para comparar as médias dos cinco grupos, possibilitando discernir se existem diferenças significativas entre elas. Esta abordagem estatística é fundamental para avaliar a influência do nível educacional na eficácia dos modelos de aprendizado por transferência. Ao analisar os dados fornecidos, foi realizada uma ANOVA que revelou resultados esclarecedores sobre a variação das médias entre os grupos conforme objetivo do experimento.

Os resultados do teste ANOVA indicarão se as diferenças entre as médias dos grupos são estatisticamente significativas, oferecendo informações valiosas sobre como o nível de ensino pode afetar a eficácia dos modelos de aprendizado por transferência. Esses resultados podem ser interpretados à luz das hipóteses estabelecidas, permitindo uma análise mais profunda do impacto do nível educacional nos resultados do estudo.

O código em *R* foi elaborado para simular dados de diferentes grupos correspondentes a diferentes níveis de ensino e realizar uma Análise de Variância (ANOVA) para comparar as médias desses grupos.

1. **Simulação de dados para cada grupo:** o código utiliza a função **rnorm** para gerar dados aleatórios seguindo uma distribuição normal para cada grupo. Os parâmetros *média* e *desvio* especificam a média e o desvio padrão de cada grupo, respectivamente. Os vetores *n* representam o tamanho de cada grupo;
2. **Combinação dos dados em um único data frame:** os dados de cada grupo são

combinados em um único *data frame* chamado ‘dados’, onde a variável *valor* contém os dados e a variável *grupo* especifica a qual *grupo* cada observação pertence.

3. **Realização do teste ANOVA:** a função **aov** é utilizada para realizar a ANOVA, com a fórmula *valor ~ grupo* indicando o teste se as médias dos grupos são iguais ou diferentes;
4. **Resumo dos resultados:** a função **summary** é aplicada ao resultado da ANOVA para fornecer um resumo dos resultados, incluindo estatísticas como a estatística *F* e os valores *p* associados, que indicam se existem diferenças significativas entre as médias dos grupos.

Com a utilização deste código, disponível em o script de ANOVA, foi possível obter informações estatísticas que permitem determinar se existem diferenças significativas entre os diversos níveis de ensino em relação à eficácia dos modelos de aprendizado por transferência no respectivo contexto educacional analisado.

Na Tabela 13 são apresentados os resultados alcançados da ANOVA.

Tabela 13 – Análise de variância (ANOVA) para grupos.

Fonte de variação	Df	Soma dos Quadrados	Média dos Quadrados	Valor F	Valor-p
Grupo	4	20814	5203	5203	<2e-16
Resíduos	235621	213702	1		

Os resultados da Tabela 13 fornecem informações sobre a variação entre os grupos (Grupos) e a variação dentro dos grupos (Resíduos), com as respectivas estatísticas de teste, como a estatística *F* e o *valor-p*.

1. Grupos:

- **Grupos:** esta fonte de variação representa a variação entre os grupos que estão sendo comparados. Em um experimento típico, os grupos podem representar diferentes tratamentos, condições ou categorias. A análise da variação entre os grupos ajuda a determinar se as médias dos grupos são estatisticamente diferentes umas das outras;
- **Resíduos:** esta fonte de variação representa a variação dentro dos grupos que não é explicada pelas diferenças entre os grupos. Refere-se à variação aleatória ou não sistemática nos dados que não pode ser atribuída aos fatores em estudo. A análise da variação residual ajuda a determinar a magnitude da variação não explicada nos dados.

2. *Df* (Graus de Liberdade:

- **Grupos:** os graus de liberdade para os grupos são calculados como o número de grupos menos 1. Nesse caso é $5 - 1 = 4$ graus de liberdade, o que significa que há 4 grupos sendo comparados;
- **Resíduos:** para calcular os graus de liberdade dos resíduos, subtrai-se o número de grupos do total de observações. No caso em questão, com um total de 235625 observações e 4 graus de liberdade para os grupos, os graus de liberdade para os resíduos são $235625 - 4 = 235621$.

3. Soma dos Quadrados:

- **Grupos:** refere-se à variação entre os grupos que estão sendo comparados. Avalia essa diferença, calculando a variação entre as médias de cada grupo e a média geral de todos os dados;
- **Resíduos:** refere-se à variação dentro dos grupos, ou seja, a variação que não pode ser explicada pelas diferenças entre os grupos. Essa variação inclui a variabilidade natural dentro de cada grupo, bem como qualquer erro de medição ou outras fontes de variação não explicadas pelas variáveis independentes (neste caso, os diferentes níveis de ensino).

4. Média dos Quadrados:

- **Grupos:** é a soma dos quadrados dividida pelos graus de liberdade. Neste caso, a soma dos quadrados que foi igual a 20814 foi dividida por 4, resultando em 5203;
- **Resíduos:** é a soma dos quadrados dividida pelos graus de liberdade. Neste caso, a soma dos quadrados de valor 213702 foi dividida por 235621, resultando em aproximadamente 1.

5. Valor *F*:

- **Grupos:** refere-se à variação nas médias dos diferentes grupos que estão sendo comparados. Esta variabilidade é calculada como a soma dos quadrados das diferenças entre a média de cada grupo e a média geral de todos os dados. A ideia é medir quanto as médias dos grupos diferem uma da outra. No caso específico, o valor de *F* é a média dos quadrados dos grupos dividido pela média dos quadrados de resíduos, ou seja, $5203 / 1 = 5203$;
- **Resíduos:** o valor de *F* não é calculado para a fonte de variação “Resíduos”

porque a estatística F é usada especificamente para testar a hipótese de que as médias dos grupos são iguais.

6. **Valor-P:**

- **Grupos:** indica a probabilidade de observar uma estatística F tão extremamente grande (ou maior) quanto a observada, assumindo que a hipótese nula (de que não há diferença entre os grupos) seja verdadeira. Um *valor-p* menor que o nível de significância (geralmente 0.05) indica que há evidências estatísticas suficientes para rejeitar a hipótese nula. Neste caso, o *valor-p* calculado foi $< 2e - 16$, o que significa que é extremamente baixo, essencialmente zero, indicando uma diferença estatisticamente significativa entre os grupos;
- **Resíduos:** o *valor-p* geralmente é utilizado como uma medida global da significância do modelo ANOVA. Este *valor-p* é obtido para testar a hipótese nula de que não há diferença significativa entre os grupos.

Para a variável “Grupo” foram obtidos 4 graus de liberdade, uma vez que havia 5 grupos distintos (“Fundamental”, “Médio”, “Superior”, “Vários”, “Não Especificado”). Os graus de liberdade para os grupos são calculados como $k - 1$, onde k é o número de grupos. Para a variável “Resíduos”, obteve-se 235621 graus de liberdade residuais, que correspondem ao total de observações menos o número de grupos.

No que se refere à “Grupo”, a soma dos quadrados foi de 20814, representando a variação explicada pelos diferentes níveis de ensino. Já para “Resíduos”, a soma dos quadrados foi de 213702, refletindo a variação não explicada dentro dos grupos.

Os quadrados médios para “Grupo” e “Resíduos” foram calculados, respectivamente, como 5203 e 1. O quadrado médio foi obtido pela divisão da soma dos quadrados pelos graus de liberdade correspondentes.

O valor F foi calculado como 5203, indicando uma grande variabilidade entre os grupos em comparação com a variabilidade dentro dos grupos, representada pelo quadrado médio dos resíduos (1). O *valor-p*, essencialmente zero, sugere uma diferença estatisticamente significativa entre os grupos.

Com um *valor-p* extremamente baixo ($< 2e - 16$), rejeita-se a hipótese nula (H_0), indicando uma diferença significativa na efetividade dos modelos de aprendizado por transferência entre os diferentes níveis de ensino.

Portanto, os resultados sugeriram fortemente que a efetividade dos modelos de

aprendizado por transferência varia significativamente entre os diferentes níveis de ensino, corroborando assim a hipótese alternativa (H1).

– Realização do Teste de Tukey

Para determinar quais pares de grupos apresentam diferenças significativas em suas médias, o Teste de Tukey foi conduzido. Essa análise adicional ajuda a entender melhor quais níveis de ensino possuem uma influência mais expressiva na eficácia dos modelos de aprendizado por transferência.

O teste de Tukey é particularmente valioso, pois oferece uma abordagem sistemática para avaliar todas as possíveis combinações entre os grupos, proporcionando uma compreensão abrangente das diferenças de média entre eles. Segundo Sheskin (2011) o teste de Tukey proporciona uma compreensão abrangente das diferenças de média, facilitando a identificação de quais grupos diferem significativamente uns dos outros, o que é crucial para análises estatísticas robustas e interpretações precisas dos dados. Isso é essencial, especialmente em experimentos com múltiplos grupos, onde pode haver variações complexas nas médias que não são evidentes apenas através da análise ANOVA.

Ao conduzir o teste de Tukey, pode-se identificar não apenas quais grupos têm médias significativamente diferentes entre si, mas também quantificar a magnitude dessas diferenças de maneira precisa e confiável. Isso permite uma interpretação mais robusta e fundamentada dos resultados, oferecendo informações valiosas sobre as relações entre os diferentes níveis de ensino e a efetividade dos modelos de aprendizado por transferência.

Portanto, ao combinar a análise ANOVA com o teste de Tukey, pode-se obter uma compreensão mais completa e detalhada das diferenças significativas entre os grupos, contribuindo assim para uma avaliação mais precisa da eficácia dos modelos de aprendizado por transferência em diferentes contextos educacionais.

Na Tabela 14 são representados os resultados do Teste de Tukey, uma técnica estatística utilizada para comparar as médias de múltiplos grupos, referente ao cenário 2 de experimentação.

A análise estatística realizada com o teste de *Tukey* para comparações múltiplas de médias entre cinco grupos distintos revelou informações significativas sobre a efetividade dos modelos de aprendizado por transferência em diferentes contextos educacionais, conforme os resultados apresentados a seguir:

Tabela 14 – Resultado do Teste de Tukey

Comparação	Diferença	Limite Inferior	Limite Superior	Valor-p ajustado
B-A	18.6725	-9.0769	46.4219	0.2955
C-A	0.1191	-30.2789	30.5171	1.0000
D-A	0.0942	-25.9371	26.1254	1.0000
E-A	0.3062	-23.6108	24.2232	0.9999995
C-B	-18.5534	-51.3870	14.2801	0.4613
D-B	-18.5784	-47.4164	10.2597	0.3355
E-B	-18.3663	-45.3112	8.5786	0.2840
D-C	-0.0249	-31.4198	31.3700	1.0000
E-C	0.1871	-29.4783	29.8525	1.0000
E-D	0.2120	-24.9599	25.3839	0.9999999

- 1. COMPARAÇÃO (Grupo B - Grupo A):** a diferença média entre o Grupo B (Ensino Médio) e o Grupo A (Ensino Fundamental) foi de 18.67, com um intervalo de confiança de (-9.08, 46.42) e um *valor-p ajustado* de 0.2955. No entanto, essa diferença não foi estatisticamente significativa;
- 2. COMPARAÇÃO (Grupo C - Grupo A):** não houve diferença estatisticamente significativa entre o Grupo C (Ensino Superior) e o Grupo A (Ensino Fundamental), com uma diferença média de 0.12 e um *valor-p ajustado* de 1.0000;
- 3. COMPARAÇÃO (Grupo D - Grupo A):** não houve diferença significativa entre o Grupo D (Vários Contextos) e o Grupo A (Ensino Fundamental), com uma diferença média de 0.09 e um *valor-p ajustado* de 1.0000;
- 4. COMPARAÇÃO (Grupo E - Grupo A):** o Grupo E (Não Especificado) também não apresentou diferença significativa em relação ao Grupo A (Ensino Fundamental), com uma diferença média de 0.31 e um *valor-p ajustado* de 0.9999995;
- 5. COMPARAÇÃO (Grupo C - Grupo B):** a diferença média entre o Grupo C (Ensino Superior) e o Grupo B (Ensino Médio) foi de -18.55, com um intervalo de confiança de (-51.39, 14.28) e um *valor-p ajustado* de 0.4613. Não houve diferença estatisticamente significativa entre esses grupos;
- 6. COMPARAÇÃO (Grupo D - Grupo B):** a diferença média entre o Grupo D (Vários Contextos) e o Grupo B (Ensino Médio) foi de -18.58, com um intervalo de confiança de (-47.42, 10.26) e um *valor-p ajustado* de 0.3355. Também não houve diferença estatisticamente significativa entre esses grupos;
- 7. COMPARAÇÃO (Grupo E - Grupo B):** a diferença média entre o Grupo E (Não Especificado) e o Grupo B (Ensino Médio) foi de -18.37, com um intervalo de confiança de (-45.31, 8.58) e um *valor-p ajustado* de 0.2840. Mais uma vez, não

houve diferença estatisticamente significativa entre esses grupos;

8. **COMPARAÇÃO (Grupo D - Grupo C):** a diferença média entre o Grupo D (Vários Contextos) e o Grupo C (Ensino Superior) foi de -0.02, com um intervalo de confiança de (-31.42, 31.37) e um *valor-p ajustado* de 1.0000. Não houve diferença estatisticamente significativa entre esses grupos;
9. **COMPARAÇÃO (Grupo E - Grupo C):** a diferença média entre o Grupo E (Não Especificado) e o Grupo C (Ensino Superior) foi de 0.19, com um intervalo de confiança de (-29.48, 29.85) e um *valor-p ajustado* de 1.0000. Novamente, não houve diferença estatisticamente significativa entre esses grupos;
10. **COMPARAÇÃO (Grupo E - Grupo D):** a diferença média entre o Grupo E (Não Especificado) e o Grupo D (Vários Contextos) foi de 0.21, com um intervalo de confiança de (-24.96, 25.38) e um *valor-p ajustado* de 0.9999999. Mais uma vez, não houve diferença estatisticamente significativa entre esses grupos.

Esses resultados sugerem que, embora existam algumas diferenças nas médias entre os grupos, muitas delas não são estatisticamente significativas. Isso indica que a efetividade dos modelos de aprendizado por transferência pode não variar de maneira relevante entre os diferentes níveis de ensino e contextos analisados. Em outras palavras, os modelos de aprendizado por transferência parecem funcionar de forma semelhante, independentemente do nível de escolaridade ou do contexto em que são aplicados.

– Conclusões sobre os Resultados do Experimento

Os resultados dos testes estatísticos, incluindo o teste ANOVA e o teste de Tukey HSD, indicaram uma clara diferença na efetividade dos modelos de aprendizado por transferência entre os diversos níveis de ensino (fundamental, médio e superior), rejeitando a hipótese nula.

Essa descoberta destacou a importância de considerar o contexto educacional ao implementar estratégias de aprendizado por transferência. A variabilidade na eficácia desses modelos em diferentes níveis de ensino sugere que as abordagens de ensino devem ser adaptadas e personalizadas para atender às necessidades específicas de cada grupo.

Portanto, pode-se concluir que a efetividade dos modelos de aprendizado por trans-

ferência varia significativamente entre os diferentes níveis de ensino, corroborando com a hipótese alternativa. Essa compreensão mais aprofundada permite uma tomada de decisão mais informada no desenvolvimento e implementação de estratégias de aprendizado, garantindo que sejam mais eficazes e relevantes para cada contexto educacional.

5.4.5.3 *Discussão dos Resultados dos Testes de Hipóteses nos Cenários 1 e 2*

Com base nos resultados dos testes estatísticos realizados para os Cenários 1 e 2 de experimentação, pode-se inferir algumas conclusões significativas sobre as métricas de desempenho e a efetividade dos modelos de aprendizado por transferência em diferentes contextos educacionais.

Para as métricas de desempenho, o teste ANOVA revelou que o *valor-p* não atingiu significância estatística ($p = 0.091$), indicando que não há evidências suficientes para rejeitar a hipótese nula de que a média das métricas de desempenho é a mesma para os três grupos. No entanto, ao realizar o teste de Tukey para comparações múltiplas entre os grupos, observou-se que não há diferenças significativas entre os grupos A e B ($p = 0.3425$) ou entre os grupos A e C ($p = 0.9499$). Entretanto, existe uma tendência à diferença entre os grupos B e C ($p = 0.0681$), sugerindo que pode haver uma discrepância nas métricas de desempenho entre esses grupos, apesar de não ser estatisticamente significativa.

No que diz respeito aos diferentes níveis de ensino, o teste ANOVA mostrou uma diferença estatisticamente significativa ($p < 0.001$) na efetividade dos modelos de aprendizado por transferência entre os grupos de ensino Fundamental, Médio e Superior. O teste de Tukey indicou que não há diferenças significativas na efetividade entre os grupos Fundamental e Médio ($p = 1.000$), Fundamental e Superior ($p = 1.000$), ou Médio e Superior ($p = 1.000$). No entanto, observou-se uma diferença estatisticamente significativa entre os grupos Fundamental e Superior ($p < 0.001$), bem como entre os grupos Médio e Superior ($p < 0.001$), sugerindo que o nível de ensino Superior pode ter uma efetividade diferente em comparação com os outros níveis de ensino.

Esses resultados destacaram a importância de analisar cuidadosamente as diferenças nas métricas de desempenho e efetividade dos modelos de aprendizado por

transferência em diferentes contextos educacionais. Embora não haja diferenças significativas nas métricas de desempenho entre os grupos analisados, há variações significativas na efetividade dos modelos entre os diferentes níveis de ensino, especialmente em relação ao Grupo 3. Essas descobertas podem ter implicações importantes para o desenvolvimento e implementação de estratégias de ensino e aprendizado em ambientes educacionais diversos.

6 Considerações Finais

A pesquisa científica é essencial para explorar e aprimorar técnicas inovadoras na educação, e este estudo investigou como a Aprendizagem por Transferência (*Transfer Learning*) pode ser aplicada em diversos contextos educacionais. Ao responder à QP1: Qual é o estado atual da pesquisa sobre o uso de *Transfer Learning* na Educação?, o estudo revelou uma ampla gama de modelos, métricas de avaliação e elementos cruciais que influenciam a eficácia e aplicabilidade dessa técnica. A revisão sistemática destacou o uso crescente da Inteligência Artificial para melhorar o processo de ensino e aprendizagem, fornecendo exemplos práticos que vão desde o diagnóstico e avaliação de distúrbios de aprendizagem até a melhoria da transferência de conhecimento em cursos online e o uso de técnicas preditivas para identificar transtornos específicos.

Essas descobertas forneceram um panorama detalhado sobre o estado atual da pesquisa na área, demonstrando a eficácia da *Transfer Learning* em diversos cenários educacionais. A análise apontou a variedade de abordagens utilizadas, as métricas mais comumente aplicadas e os desafios enfrentados para alcançar resultados satisfatórios, contribuindo assim para um avanço significativo no campo, conforme destacado na QP1.

Este estudo representa um esforço significativo para analisar a literatura sobre Aprendizagem por Transferência na Educação, abordando duas fases principais: uma Revisão Sistemática da Literatura (RSL) e uma Meta-análise. A RSL foi conduzida de acordo com um protocolo rigoroso, o que permitiu a identificação de estudos relevantes e a extração de dados para responder às questões de pesquisa definidas no protocolo. A meta-análise complementou os resultados da RSL com análises estatísticas robustas, oferecendo evidências sólidas que aprofundam a compreensão do tema. O trabalho expande o conhecimento atual e oferece orientações valiosas para futuras pesquisas e práticas educacionais.

Na QP2: Quais são os principais domínios educacionais em que *Transfer Lear-*

ning tem sido aplicado com sucesso?, foram identificados os principais domínios educacionais, como matemática e línguas, onde o *Transfer Learning* foi aplicado com sucesso. A Revisão Sistemática da Literatura (RSL) permitiu identificar estudos que demonstraram melhorias no desempenho dos alunos em diversas disciplinas. Além disso, a meta-análise ofereceu uma visão detalhada sobre o impacto do *Transfer Learning* em tarefas educacionais específicas, como classificação de textos e recomendação de conteúdo, abordando a QP4: Como o *Transfer Learning* tem sido usado para melhorar o desempenho dos alunos em tarefas específicas?.

Ao examinar os dados, a QP3: Quais são os conjuntos de dados mais frequentemente usados em pesquisas de *Transfer Learning* na Educação? foi respondida ao identificar os conjuntos de dados mais utilizados em pesquisas de *Transfer Learning* na Educação. A análise destacou a diversidade de dados usados, suas características e a forma como essas informações influenciam os resultados.

Os desafios técnicos e práticos mencionados na QP5: Quais são os desafios técnicos e práticos associados à implementação bem-sucedida de *Transfer Learning* em contextos educacionais? foram evidentes em muitos dos estudos analisados. A dificuldade de adaptação dos modelos a novos contextos educacionais, assim como a falta de dados de alta qualidade, foram questões recorrentes, exigindo soluções criativas para garantir a implementação eficaz.

A análise também abordou preocupações éticas importantes, respondendo à QP6: Quais são as limitações e as preocupações éticas relacionadas ao uso de *Transfer Learning* na Educação? Questões como viés algorítmico e privacidade dos dados foram identificadas como limitações significativas na pesquisa de *Transfer Learning*, com recomendações para práticas mais transparentes e equitativas.

Por fim, na QP7: Quais são as lacunas na pesquisa atual e as áreas que exigem mais investigação em relação ao uso de *Transfer Learning* na Educação?, as lacunas e oportunidades de pesquisa foram destacadas. A falta de validação das soluções com usuários finais e a necessidade de explorar o *Transfer Learning* em diferentes ambientes educacionais mostram que ainda há muito a ser investigado. Essas lacunas sugerem caminhos promissores para futuras pesquisas, especialmente em termos de

generalização e personalização dos modelos para contextos educacionais variados.

A revisão sistemática e a meta-análise proporcionaram uma visão abrangente do estado atual da pesquisa nesta área. Ao examinar uma ampla gama de estudos, essas metodologias permitiram identificar tendências e padrões, além de oferecer perspectivas sobre a aplicação da Aprendizagem por Transferência na Educação. Foi possível destacar práticas recentes e áreas de sucesso, bem como oportunidades para avanços futuros.

Os resultados revelaram a eficácia, as limitações e as potenciais aplicações da Aprendizagem por Transferência em contextos educacionais diversos. A diversidade de abordagens, técnicas e métricas de desempenho utilizadas ficou evidente. A revisão e a meta-análise ressaltaram a importância da validação rigorosa e da consideração dos usuários finais na avaliação das soluções, garantindo a eficácia, a usabilidade e a relevância prática dos modelos.

A análise abordou os objetivos dos estudos, os modelos implementados e os desempenhos alcançados. Observou-se que os procedimentos de validação e as métricas de desempenho variam entre os estudos, refletindo metodologias adaptadas a diferentes cenários. Embora muitos artigos não detalhassem o desempenho dos modelos, aqueles que o fizeram geralmente relataram melhorias significativas, evidenciando a eficácia das abordagens propostas. Essa constatação sublinha a importância de uma análise cuidadosa do desempenho dos modelos no contexto da Aprendizagem por Transferência na Educação.

Refletir sobre as lições aprendidas é crucial para o crescimento pessoal e profissional, consolidando conquistas passadas e orientando decisões futuras. No âmbito do *Transfer Learning*, que aplica conhecimento adquirido a problemas semelhantes, foram identificadas áreas de melhoria durante este trabalho, a saber:

1. Variedade de estudos e abordagens: a RSL pôde revelar uma ampla variedade de estudos e abordagens relacionadas ao *Transfer Learning* na Educação, destacando diferentes contextos de aplicação, modelos utilizados e métricas de avaliação;

2. Importância da estratégia de busca: a definição de uma estratégia de

busca adequada foi fundamental para garantir a identificação abrangente e precisa de estudos relevantes. Lições aprendidas incluem a necessidade de ajustar e aprimorar a estratégia de busca ao longo do processo para garantir a inclusão de todos os estudos pertinentes;

3. Consideração de diferentes tipos de publicações: a inclusão de diferentes tipos de publicações, como artigos de periódicos, conferências e teses, pode enriquecer a revisão sistemática, fornecendo uma visão mais abrangente do campo de estudo;

4. Validação dos resultados: a validação dos resultados obtidos na RSL por meio da meta-análise pode ajudar a quantificar e sintetizar os efeitos do *Transfer Learning* na Educação, fornecendo *insights* adicionais sobre a eficácia e a generalização das abordagens investigadas;

5. Identificação de lacunas e oportunidades de pesquisa: através da RSL e da meta-análise foi possível destacar as lacunas na literatura existente e identificar oportunidades de pesquisas futuras, como áreas sub-representadas, temas emergentes e questões não abordadas.

Essas lições aprendidas proporcionam uma visão mais abrangente e completa do estado atual da pesquisa sobre *Transfer Learning* na Educação, contribuindo para o avanço do campo. Considerando as limitações e observações derivadas da aplicação dessa técnica, é crucial abordar restrições práticas, como a qualidade e representatividade dos dados e a generalização das soluções propostas.

As ameaças à validade do estudo sobre o uso de Aprendizagem por Transferência (*Transfer Learning*) na Educação podem ser classificadas em várias dimensões. Um ponto crucial a ser considerado é a quantidade e a qualidade dos dados utilizados. A pesquisa de Ding *et al.* (2019) destacou a dependência de dados de *clickstream*, limitando a generalização da abordagem a ambientes de aprendizado *online* que não coletam esses dados. Essa limitação sugere que as conclusões tiradas a partir de determinados contextos podem não se aplicar amplamente a outros cenários educacionais.

Outro aspecto a ser considerado é a falta de consideração de informações contex-

tuais latentes e objetivos de aprendizado. Essa ausência pode restringir a capacidade da solução de fornecer *insights* realmente úteis para educadores e alunos, comprometendo a relevância prática das abordagens analisadas. Além disso, a ausência de validação com usuários finais pode resultar em uma compreensão limitada da eficácia da solução em um ambiente real de aprendizado *online*. Essa limitação é especialmente preocupante, pois a aceitação e a aplicabilidade das soluções propostas dependem fortemente da percepção e experiência dos educadores e alunos que as utilizarão.

Os desafios relacionados à interpretabilidade dos modelos de *transfer learning* também foram mencionados, evidenciando a necessidade de tornar esses modelos mais transparentes e compreensíveis. Para que educadores e alunos possam confiar nos resultados gerados, a clareza na operação dos modelos é essencial. Além disso, o viés nos dados de treinamento, conforme destacado por Mor e Dardeck (2021), pode afetar as generalizações dos modelos, levando a resultados tendenciosos que não refletem com precisão a diversidade dos alunos e seus contextos.

Outro fator que merece atenção é a heterogeneidade dos estudos incluídos na revisão sistemática. As diferenças em contexto, métodos, métricas de avaliação e resultados podem dificultar a síntese e a comparação dos dados. Essa diversidade implica que as conclusões tiradas de um conjunto de estudos podem não ser aplicáveis a outros, gerando incertezas sobre a eficácia das abordagens de *Transfer Learning* em diferentes situações educacionais.

Além disso, a disponibilidade limitada de dados em algumas áreas, ou a falta de transparência nos relatórios dos estudos primários, pode dificultar a análise e a interpretação dos resultados, levando a possíveis omissões de informações cruciais. A eficácia dos modelos de *transfer learning* também pode variar em diferentes contextos educacionais, conforme apontado por Jensen *et al.* (2021). Essa variação pode ser atribuída a fatores como diferenças culturais, estruturas curriculares e estilos de aprendizado dos alunos, que influenciam diretamente a aplicabilidade dos modelos analisados.

Ao reconhecer e abordar essas limitações, é possível garantir a confiabilidade e

a validade dos resultados obtidos. Essas reflexões oferecem direções valiosas para futuras pesquisas e práticas educacionais, destacando a importância de continuar a explorar abordagens adaptáveis e personalizadas que atendam às necessidades diversas dos alunos e dos ambientes de aprendizagem em que estão inseridos.

REFERÊNCIAS

- ALQAHTANI, N. D.; ALZAHRANI, B.; RAMZAN, M. S. Deep learning applications for dyslexia prediction. **Applied Sciences**, MDPI, v. 13, n. 5, p. 2804, 2023.
- ARNOLD, K. E.; PISTILLI, M. D. Feedback for learning: Closing the assessment loop. **Assessment & Evaluation in Higher Education**, Taylor & Francis, v. 37, n. 2, p. 219–232, 2012.
- BARTLETT, M. S. Properties of sufficiency and statistical tests. **Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences**, v. 160, n. 901, p. 268–282, 1937.
- BASSOK, M.; HOLYOAK, K. J. Interdomain transfer between isomorphic topics in algebra and physics. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, American Psychological Association, v. 15, n. 1, p. 153, 1989.
- BENEDETTO, L.; CREMONESI, P.; CAINES, A.; BUTTERY, P.; CAPPELLI, A.; GIUSSANI, A.; TURRIN, R. A survey on recent approaches to question difficulty estimation from text. **ACM Computing Surveys**, ACM New York, NY, v. 55, n. 9, p. 1–37, 2023.
- BENSON, G. **High School Education: Enhancing Preparation for Higher Education and the Workforce**. [S.l.]: Springer, 2018.
- BERK, L. E. **Child Development**. 9th. ed. [S.l.]: Pearson, 2018. 278–319 p.
- BIAN, C.; ZHANG, Y.; YANG, F.; BI, W.; LU, W. Spontaneous facial expression database for academic emotion inference in online learning. **IET Computer Vision**, Wiley Online Library, v. 13, n. 3, p. 329–337, 2019.
- BONTHU, S.; SREE, S. R.; PRASAD, M. K. Improving the performance of automatic short answer grading using transfer learning and augmentation. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 123, p. 106292, 2023.
- BORENSTEIN, M.; HEDGES, L. V.; HIGGINS, J. P.; ROTHSTEIN, H. R. **Introduction to Meta-Analysis**. [S.l.]: John Wiley & Sons, 2009.
- BORENSTEIN, M.; HEDGES, L. V.; HIGGINS, J. P. T.; ROTHSTEIN, H. R. **Introduction to Meta-Analysis**. [S.l.]: John Wiley & Sons, 2011.
- BRIZOLA, J.; FANTIN, N. Revisão da literatura e revisão sistemática da literatura. **Revista de Educação do Vale do Arinos-RELVA**, v. 3, n. 2, 2016.
- BROWN, M.; FORSYTHE, A. Robust tests for the equality of variances. **Journal of the American Statistical Association**, v. 69, n. 346, p. 364–367, 1974.
- BRUSILOVSKY, P.; HORNBÆK, S. **Adaptive Educational Technologies for Literacy Instruction**. [S.l.]: Springer, 2016.
- BUZAN, T.; BUZAN, B. **The Mind Map Book: Unlock your creativity, boost your memory, change your life**. London: BBC Active, 2006.

- CAMPBELL, M. S. P. **Formulation of Discrete Event Simulation Models for Economic Impact Analysis**. Hershey, PA: Idea Group Inc (IGI), 2023.
- CASTELLS, M. **The Rise of the Network Society: The Information Age: Economy, Society and Culture**. [S.l.]: Wiley-Blackwell, 2000. v. 1.
- CAVALIERE, A. M. Tempo de escola e qualidade na educação pública. **Educação & Sociedade**, SciELO Brasil, v. 28, p. 1015–1035, 2007.
- CHEN, L.; CHEN, G.; LIN, J. Artificial intelligence in education: A review. **IEEE Access**, IEEE, v. 8, p. 75264–75278, 2020.
- COOPER, G.; SWELLER, J. Effects of schema acquisition and rule automation on mathematical problem-solving transfer. **Journal of educational psychology**, American Psychological Association, v. 79, n. 4, p. 347, 1987.
- DEEKS, J. J.; HIGGINS, J. P.; ALTMAN, D. G. Chapter 10: Analysing data and undertaking meta-analyses. **Cochrane Handbook for Systematic Reviews of Interventions**, John Wiley & Sons, p. 241–284, 2019.
- DELANEY, V.; BHATIA, J. Classification in math class: Using convolutional neural networks to categorize student cognitive demand. In: **CEUR Workshop Proceedings**. [S.l.: s.n.], 2021.
- DERMEVAL, D.; COELHO, J. A. d. M.; BITTENCOURT, I. I. Mapeamento sistemático e revisão sistemática da literatura em informática na educação. **JAQUES, Patrícia Augustin; SIQUEIRA, Sean; BITTENCOURT, Ig; PIMENTEL, Mariano.(Org.) Metodologia de Pesquisa Científica em Informática na Educação: Abordagem Quantitativa**. Porto Alegre: SBC, 2020.
- DERSIMONIAN, R.; KACKER, R. Random-effects model for meta-analysis of clinical trials: an update. **Contemporary clinical trials**, Elsevier, v. 28, n. 2, p. 105–114, 2007.
- DING, M.; WANG, Y.; HEMBERG, E.; O'REILLY, U.-M. Transfer learning using representation learning in massive open online courses. In: **Proceedings of the 9th International Conference on Learning Analytics & Knowledge**. [S.l.: s.n.], 2019. p. 145–154.
- DONATO, H.; DONATO, M. Etapas na condução de uma revisão sistemática. **Acta Médica Portuguesa**, v. 32, n. 3, p. 227–235, 2019.
- EGGER, M.; SMITH, G. D.; SCHNEIDER, M.; MINDER, C. Bias in meta-analysis detected by a simple, graphical test. **BMJ**, BMJ Publishing Group Ltd, v. 315, n. 7109, p. 629–634, 1997.
- ESTEBAN, M. T. Educação popular: desafio à democratização da escola pública. **Cadernos Cedes**, SciELO Brasil, v. 27, p. 9–17, 2007.
- FAN, R.; ALWAN, A. Draft: A novel framework to reduce domain shifting in self-supervised learning and its application to children's asr. **arXiv preprint arXiv:2206.07931**, 2022.
- FAWCETT, T. An introduction to roc analysis. **Pattern Recognition Letters**, Elsevier, v. 27, n. 8, p. 861–874, 2006.

FISHER, R. A. **Statistical Methods for Research Workers**. Edinburgh, Scotland: Oliver and Boyd, 1925.

FUCHS, S. C. P. C.; PAIM, B. S. Revisão sistemática de estudos observacionais com metanálise. **Revista HCPA. Porto Alegre. Vol. 30, n. 3 (2010), p. 294-301**, 2010.

FUNG, Y.-C.; KWOK, J. C.-W.; LEE, L.-K.; CHUI, K. T.; U, L. H. Automatic question generation system for english reading comprehension. In: SPRINGER. **Technology in Education. Innovations for Online Teaching and Learning: 5th International Conference, ICTE 2020, Macau, China, August 19-22, 2020, Revised Selected Papers 5**. [S.l.], 2020. p. 136–146.

GELMAN, A.; HILL, J. **Data Analysis Using Regression and Multilevel/Hierarchical Models**. [S.l.]: Cambridge University Press, 2006.

GÉRON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. [S.l.]: O'Reilly Media, 2019.

GILLANI, N.; EYNON, R.; CHIABAUT, C.; FINKEL, K. Unpacking the “black box” of ai in education. **Educational Technology & Society**, JSTOR, v. 26, n. 1, p. 99–111, 2023.

GOMES, L. F.; PEREIRA, M. S.; SANTOS, R. M. Strategic application of educational programs to enhance teaching and learning processes. **International Journal of Education and Development using ICT**, v. 14, n. 1, p. 56–72, 2019.

GONÇALVES, H. d. A.; NASCIMENTO, M. B. d. C.; NASCIMENTO, K. C. S. Revisão sistemática e metanálise: níveis de evidência e validade científica. **Revista Eletrônica Debates em Educação Científica e Tecnológica**, Instituto Federal do Espírito Santo (IFES), 2015.

GRANT, M. J.; BOOTH, A. A typology of reviews: an analysis of 14 review types and associated methodologies. **Health Information & Libraries Journal**, Wiley Online Library, v. 26, n. 2, p. 91–108, 2009.

GREENHALGH, T. Assessing the quality of published papers. **BMJ: British Medical Journal**, British Medical Journal Publishing Group, v. 315, n. 7103, p. 305–308, 1997.

GUIMARÃES, C. A.; NUNES, I.; PIRES, A. K.; ALENCAR, E. E. da S. A produção de learning analytics e predição de desempenho acadêmico por pesquisadores brasileiros: Uma revisão sistemática da literatura. In: SBC. **Anais do V Congresso sobre Tecnologias na Educação**. [S.l.], 2020. p. 306–315.

HAMIDA, S.; GANNOUR, O. E.; CHERRADI, B.; OUAJJI, H.; RAIHANI, A. Handwritten computer science words vocabulary recognition using concatenated convolutional neural networks. **Multimedia Tools and Applications**, Springer, v. 82, n. 15, p. 23091–23117, 2023.

HARDALOV, M.; MIHAYLOV, T.; ZLATKOVA, D.; DINKOV, Y.; KOYCHEV, I.; NAKOV, P. Exams: A multi-subject high school examinations dataset for cross-lingual and multilingual question answering. **arXiv preprint arXiv:2011.03080**, 2020.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: Data mining, inference, and prediction**. [S.l.]: Springer Science & Business Media, 2009.

HATTIE, J. Visible learning: A synthesis of over 800 meta-analyses relating to achievement. **Routledge**, 2009.

HIGGINS, J. P.; GREEN, S. **Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0**. [S.l.]: John Wiley & Sons, 2011.

HIGGINS, J. P.; THOMPSON, S. G. Quantifying heterogeneity in a meta-analysis. **Statistics in medicine**, Wiley Online Library, v. 21, n. 11, p. 1539–1558, 2002.

HIGGINS, J. P.; THOMPSON, S. G.; DEEKS, J. J.; ALTMAN, D. G. Measuring inconsistency in meta-analyses. **BMJ (Clinical Research Ed.)**, BMJ Publishing Group Ltd, v. 327, n. 7414, p. 557–560, 2003.

HIGGINS, J. P. T.; GREEN, S. **Cochrane Handbook for Systematic Reviews of Interventions**. 2nd. ed. Wiley-Blackwell, 2019. Disponível em: <<https://training.cochrane.org/handbook>>. Disponível em: <<https://training.cochrane.org/handbook>>.

HIRAKATA, V. N.; MANCUSO, A. C. B.; CASTRO, S. M. de J. Teste de hipóteses: perguntas que você sempre quis fazer, mas nunca teve coragem. **Clinical and Biomedical Research**, v. 39, n. 2, 2019.

HOLT, R. D.; SMITH, J. K. **Statistical Methods for the Analysis of Variance**. Cambridge, UK: Cambridge University Press, 2010.

HU, M.; WEI, Y.; LI, M.; YAO, H.; DENG, W.; TONG, M.; LIU, Q. Bimodal learning engagement recognition from videos in the classroom. **Sensors**, MDPI, v. 22, n. 16, p. 5932, 2022.

HUANG, H.; GUO, S. Siamese-ernie: Topic relevance check for english writing. In: **2021 2nd International Conference on Computers, Information Processing and Advanced Education**. [S.l.: s.n.], 2021. p. 994–997.

HUNTER, J. E.; SCHMIDT, F. L. **Methods of Meta-Analysis: Correcting Error and Bias in Research Findings**. [S.l.]: SAGE Publications, 2004.

IKRAM, S.; AHMAD, H.; MAHMOOD, N.; FAISAL, C. N.; ABBAS, Q.; QURESHI, I.; HUSSAIN, A. Recognition of student engagement state in a classroom environment using deep and efficient transfer learning algorithm. **Applied Sciences**, MDPI, v. 13, n. 15, p. 8637, 2023.

INTHOUT, J.; IOANNIDIS, J. P.; BORM, G. F. Small studies are more heterogeneous than large ones: a meta-meta-analysis. **Journal of Clinical Epidemiology**, Elsevier, v. 69, p. 238–246, 2016.

JENSEN, E.; PUGH, S. L.; D’MELLO, S. K. A deep transfer learning approach to modeling teacher discourse in the classroom. In: **LAK21: 11th international learning analytics and knowledge conference**. [S.l.: s.n.], 2021. p. 302–312.

KAHAWANUGODA, A.; GNANARATHNA, K.; MEEGODA, N.; MONARAWILA, R.; SAMARASINGHE, P.; LINDAMULAGE, A. G. Development of low resource machine learning models for child cognitive ability assessments. In:

IEEE. **2022 4th International Conference on Advancements in Computing (ICAC)**. [S.l.], 2022. p. 72–77.

KARIMAH, S. N.; HASEGAWA, S. Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods. **Smart Learning Environments**, Springer, v. 9, n. 1, p. 31, 2022.

KITCHENHAM, B.; CHARTERS, S. Guidelines for performing systematic literature reviews in software engineering. Citeseer, 2007.

KUNHOTH, J.; MAADEED, S. A.; SALEH, M.; AKBARI, Y. Cnn feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children. **Expert Systems with Applications**, Elsevier, p. 120740, 2023.

LEHRER, R.; LITTLEFIELD, J. Relationships among cognitive components in logo learning and transfer. **Journal of Educational Psychology**, American Psychological Association, v. 85, n. 2, p. 317, 1993.

LEVENE, H. Robust tests for equality of variances. **Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling**, p. 278–292, 1960.

LEVENE, H. Robust tests for equality of variances. **Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling**, p. 278–292, 1960.

LIKERT, R. A technique for the measurement of attitudes. **Archives of Psychology**, Columbia University Press, New York, v. 140, p. 1–55, 1932.

LINDGREN, R.; MORPHEW, J. W.; KANG, J.; PLANEY, J.; MESTRE, J. P. Learning and transfer effects of embodied simulations targeting crosscutting concepts in science. **Journal of Educational Psychology**, American Psychological Association, v. 114, n. 3, p. 462, 2022.

LU, X. Deep learning based emotion recognition and visualization of figural representation. **Frontiers in psychology**, Frontiers Media SA, v. 12, p. 818833, 2022.

LUGOBONI, L. F. Identificação de uma lacuna de pesquisa. **Revista Liceu On-Line**, v. 7, n. 2, p. 1–5, 2017.

MARQUES, L. D. *et al.* Modelos dinâmicos com dados em painel: revisão de literatura. **Centro de estudos Macroeconômicos e Previsão, faculdade de Economia do Porto**, v. 30, p. 37, 2000.

MOHAMED, B. A.; ABDELHAKIM, B. A.; YOUNESS, S. A deep learning model for an intelligent chat bot system: an application to e-learning domain. In: SPRINGER. **Innovations in Smart Cities Applications Volume 4: The Proceedings of the 5th International Conference on Smart City Applications**. [S.l.], 2021. p. 165–179.

MOR, N. S.; DARDECK, K. Applying a convolutional neural network to screen for specific learning disorder. **Learning Disabilities: A Contemporary Journal**, ERIC, v. 19, n. 2, p. 161–169, 2021.

MORTON, S. C.; SCOTT, M. A.; AL. *et.* The forest plot in meta-analysis: A primer. **Statistical Medicine**, v. 28, n. 29, p. 3796–3812, 2009.

- MUNOZ, S. I. S.; TAKAYANAGUI, A. M. M.; SANTOS, C. B. d. *et al.* Revisão sistemática de literatura e metanálise: noções básicas sobre seu desenho, interpretação e aplicação na área da saúde. In: SCIELO BRASIL. **Proceedings of the 8. Brazilian Nursing Communication Symposium**. [S.l.], 2002.
- OLIVEIRA, J. A.; SILVA, M. R.; SANTOS, P. L. Critérios de qualidade em publicações científicas. **Revista Acadêmica de Pesquisas**, v. 15, n. 2, p. 123–145, 2021.
- ORDÓÑEZ, C. L.; CARLO, M. S.; SNOW, C. E.; MCLAUGHLIN, B. Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? **Journal of educational psychology**, American Psychological Association, v. 94, n. 4, p. 719, 2002.
- PAN, S. J.; YANG, Q. A survey on transfer learning. **IEEE Transactions on knowledge and data engineering**, IEEE, v. 22, n. 10, p. 1345–1359, 2009.
- PAN, S. J.; YANG, Q. A survey on transfer learning. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 22, n. 10, p. 1345–1359, 2010.
- PASSOS, C. L.; NACARATO, A.; FIORENTINI, D.; MISKULIN, R. G.; GRANDO, R. C.; GAMA, R.; MEGID, M. A.; FREITAS, M. T.; MELO, M. V. de. Desenvolvimento profissional do professor que ensina matemática: uma meta-análise de estudos brasileiros. **Quadrante**, v. 15, n. 1&2, p. 193–219, 2006.
- PERSEGHIN, E.; FORESTI, G. L. A shallow system prototype for violent action detection in italian public schools. **Information**, MDPI, v. 14, n. 4, p. 240, 2023.
- POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness & correlation. **Journal of Machine Learning Technologies**, Microtome Publishing, v. 2, n. 1, p. 37–63, 2011.
- QUEIROZ, A. C. M.; TORI, R.; NASCIMENTO, A. M. Realidade virtual na educação: panorama dos grupos de pesquisa no brasil. In: **Congresso Da Sociedade Brasileira De Computação**. [S.l.: s.n.], 2017. v. 4, p. 203–212.
- RAUDENBUSH, S. W.; BRYK, A. S. **Hierarchical Linear Models: Applications and Data Analysis Methods**. [S.l.]: Sage Publications, 2002.
- RILEY, R. D.; HIGGINS, J. P.; DEEKS, J. J. Interpretation of random effects meta-analyses. **BMJ**, BMJ Publishing Group Ltd, v. 342, p. d549, 2011.
- ROCHA, S. P. V. Tornar-se quem se é: educação como formação, educação como transformação. **ACTAS**, v. 3, 2016.
- ROSLI, M. S. A. B.; ISA, I. S.; RAMLAN, S. A.; SULAIMAN, S. N.; MARUZUKI, M. I. F. Development of cnn transfer learning for dyslexia handwriting recognition. In: IEEE. **2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)**. [S.l.], 2021. p. 194–199.
- SAADNA, Y.; BOUDHIR, A. A.; AHMED, M. B. An analysis of resnet50 model and rmsprop optimizer for education platform using an intelligent chatbot system. In: SPRINGER. **Networking, Intelligent Systems and Security: Proceedings of NISS 2021**. [S.l.], 2022. p. 577–590.

SABUNCUOĞLU, A.; SEZGIN, T. M. Developing affordable tangible programming education applications using mobile vision. In: IEEE. **2021 29th Signal Processing and Communications Applications Conference (SIU)**. [S.l.], 2021. p. 1–4.

SANTOS, E.; CUNHA, M. Interpretação crítica dos resultados estatísticos de uma meta-análise: estratégias metodológicas. **Millenium-Journal of Education, Technologies, and Health**, n. 44, p. 85–98, 2016.

SHARMILA, C.; SHANTHI, N.; SANTHIYA, S.; SARAN, E.; RAKESH, K. S.; SRUTHI, R. An automated system for the early detection of dysgraphia using deep learning algorithms. In: IEEE. **2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)**. [S.l.], 2023. p. 251–257.

SHEKIN, D. J. **Handbook of Parametric and Nonparametric Statistical Procedures**. 5th. ed. [S.l.]: CRC Press, 2011. 570–592 p.

SIEMENS, G.; DAWSON, S.; LYNCH, G.; BAKER, R. Learning analytics: The emergence of a discipline. **American Behavioral Scientist**, SAGE Publications Sage CA: Los Angeles, CA, v. 57, n. 10, p. 1380–1400, 2013.

SILVA, A. B. The impact of digital technologies on education: A review of current evidence. **Journal of Educational Technology**, v. 15, n. 2, p. 23–45, 2020.

SILVA, F. L. D.; COSTA, A. H. R. A survey on transfer learning for multiagent reinforcement learning systems. **Journal of Artificial Intelligence Research**, v. 64, p. 645–703, 2019.

SMITH, J. **Educational Contexts and Transfer of Learning**. New York: Educational Press, 2023. ISBN 978-1-2345-6789-0.

SMITH, J.; JOHNSON, A. Meta-analysis: A comprehensive statistical tool. **Journal of Educational Research**, v. 25, n. 3, p. 123–135, 2019.

SMITH, J. A. A natureza subjetiva da qualidade em publicações acadêmicas: Uma revisão crítica. **Revista de Estudos Acadêmicos**, v. 12, n. 3, p. 45–67, 2020.

SNIJDERS, T. A.; BOSKER, R. J. **Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling**. [S.l.]: Sage, 2012.

SOKOLOVA, M.; LAPALME, G. A systematic analysis of performance measures for classification tasks. **Information Processing & Management**, Elsevier, v. 45, n. 4, p. 427–437, 2009.

SOUSA, M. J.; MOURA, R. Innovation in the education sector: Technology and creativity. **Educational Technology & Society**, v. 21, n. 4, p. 1–14, 2018.

SPANHOL, F.; BENETTI, K.; GIGLIO, K.; FREIRE, P. O estado da arte da educação a distância: uma meta-análise da contribuição da ufsc. In: **CIAED CONGRESSO INTERNACIONAL ABED DE EDUCAÇÃO A DISTÂNCIA**. [S.l.: s.n.], 2010. v. 16.

StataCorp. **Stata Statistical Software: Release 17**. College Station, TX, 2021. Disponível em: <<https://www.stata.com>>.

SWAMY, V.; MARRAS, M.; KÄSER, T. Meta transfer learning for early success prediction in moocs. In: **Proceedings of the Ninth ACM Conference on Learning@ Scale**. [S.l.: s.n.], 2022. p. 121–132.

TAYLOR, M. E.; STONE, P. Transfer learning for reinforcement learning domains: A survey. **Journal of Machine Learning Research**, v. 10, n. 7, 2009.

TERAN-QUEZADA, A.; LOPEZ-CABRERA, V.; RANGEL, J. C.; SANCHEZ-GALAN, J. E. Hand gesture recognition with convnets for school-aged children to learn basic arithmetic operations. In: IEEE. **2022 IEEE 40th Central America and Panama Convention (CONCAPAN)**. [S.l.], 2022. p. 1–6.

TORREY, L.; SHAVLIK, J. Transfer learning. In: **Handbook of research on machine learning applications and trends: algorithms, methods, and techniques**. [S.l.]: IGI global, 2010. p. 242–264.

VACHEV, K.; HARDALOV, M.; KARADZHOV, G.; GEORGIEV, G.; KOYCHEV, I.; NAKOV, P. Generating answer candidates for quizzes and answer-aware question generators. **arXiv preprint arXiv:2108.12898**, 2021.

VASCONCELOS, T. A importância da educação na construção da cidadania. 2007.

VIECHTBAUER, W. Conducting meta-analyses in r with the metafor package. **Journal of Statistical Software**, Foundation for Open Access Statistics, v. 36, n. 3, p. 1–48, 2010.

WANG, R.; SHI, Z. Personalized online education learning strategies based on transfer learning emotion classification model. **Security and Communication Networks**, Hindawi Limited, v. 2021, p. 1–11, 2021.

WEISS, K.; KHOSHGOFTAAR, T. M.; WANG, D. A survey of transfer learning. **Journal of Big Data**, Springer, v. 3, n. 1, p. 9, 2016.

XIAO, Y.; WANG, T.; SUN, X.; LI, Y.; SONG, Y.; CUI, J.; JIA, Q.; LIU, C.; GEHRINGER, E. F. Modeling review helpfulness with augmented transformer neural networks. In: IEEE. **2022 IEEE 16th International Conference on Semantic Computing (ICSC)**. [S.l.], 2022. p. 83–90.

YANG, C.; HUANG, R.; YU, X.; PENG, R. Math word problem solver based on text-to-text transformer model. In: IEEE. **2021 IEEE International Conference on Engineering, Technology & Education (TALE)**. [S.l.], 2021. p. 818–822.

YING, W.; ZHANG, Y.; HUANG, J.; YANG, Q. Transfer learning via learning to transfer. In: PMLR. **International Conference on Machine Learning**. [S.l.], 2018. p. 5085–5094.

ZHUANG, F.; QI, Z.; DUAN, K.; XI, D.; ZHU, Y.; ZHU, H.; XIONG, H.; HE, Q. A comprehensive survey on transfer learning. **Proceedings of the IEEE**, IEEE, v. 109, n. 1, p. 43–76, 2020.

APÊNDICES

A ARTIGOS SELECIONADOS NA RSL

TÍTULO	CRITÉRIO DE SELEÇÃO
Learning and Transfer Effects of Embodied Simulations Targeting Crosscutting Concepts in Science	Apresentar uma solução usando aprendizagem por transferência.
Generating Answer Candidates for Quizzes and Answer-Aware Question Generators	Apresentar uma solução considerando o contexto educacional.
Bimodal Learning Engagement Recognition from Videos in the Classroom	Apresentar uma solução usando aprendizagem por transferência.
A Survey on Recent Approaches to Question Difficulty Estimation from Text	Apresentar uma solução considerando o contexto educacional.
Recognition of Student Engagement State in a Classroom Environment Using Deep and Efficient Transfer Learning Algorithm	Apresentar uma solução usando aprendizagem por transferência.
Unpacking the “Black Box” of AI in Education	Apresentar uma solução considerando o contexto educacional.
Siamese-ERNIE: Topic Relevance Check for English Writing	Apresentar uma solução considerando o contexto educacional.
Transfer Learning using Representation Learning in Massive Open Online Courses	Apresentar uma solução considerando o contexto educacional.
CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children	Apresentar uma solução usando aprendizagem por transferência.
Development of Low Resource Machine Learning Models for Child Cognitive Ability Assessments	Apresentar uma solução considerando o contexto educacional.
Meta Transfer Learning for Early Success Prediction in MOOCs	Apresentar uma solução usando aprendizagem por transferência.
Classification in Math Class: Using Convolutional Neural Networks to Categorize Student Cognitive Demand	Apresentar uma solução considerando o contexto educacional.
A Deep Transfer Learning Approach to Modeling Teacher Discourse in the Classroom	Apresentar uma solução usando aprendizagem por transferência.
Math Word Problem Solver Based on Text-to-Text Transformer Model	Apresentar uma solução usando aprendizagem por transferência.
Continua na próxima página	

TÍTULO	CRITÉRIO DE SELEÇÃO
Spontaneous facial expression database for academic emotion inference in online learning	Apresentar uma solução considerando o contexto educacional.
Interdomain Transfer Between Isomorphic Topics in Algebra and Physics	Apresentar uma solução considerando o contexto educacional.
Depth and Breadth of Vocabulary in Two Languages: Which Vocabulary Skills Transfer?	Apresentar uma solução considerando o contexto educacional.
Relationships Among Cognitive Components in Logo Learning and Transfer	Apresentar uma solução usando aprendizagem por transferência.
EXAMS: A Multi-Subject High School Examinations Dataset for Cross-Lingual and Multilingual Question Answering	Apresentar uma solução considerando o contexto educacional.
Development of CNN Transfer Learning for Dyslexia Handwriting Recognition	Apresentar uma solução usando aprendizagem por transferência.
DRAFT: A Novel Framework to Reduce Domain Shifting in Self-supervised Learning and Its Application to Children's ASR	Apresentar uma solução usando aprendizagem por transferência.
Modeling review helpfulness with augmented transformer neural networks	Apresentar uma solução considerando o contexto educacional.
Deep Learning Based Emotion Recognition and Visualization of Figural Representation	Apresentar uma solução considerando o contexto educacional.
Hand Gesture Recognition with ConvNets for School-Aged Children to Learn Basic Arithmetic Operations	Apresentar uma solução usando aprendizagem por transferência.
Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods	Apresentar uma solução considerando o contexto educacional.
Deep Learning Applications for Dyslexia Prediction	Apresentar uma solução considerando o contexto educacional.
An Automated System for the Early Detection of Dysgraphia using Deep Learning Algorithms	Apresentar uma solução usando aprendizagem por transferência.
Applying a Convolutional Neural Network to Screen for Specific Learning Disorder	Apresentar uma solução considerando o contexto educacional.
A Shallow System Prototype for Violent Action Detection in Italian Public Schools	Apresentar uma solução considerando o contexto educacional.
Continua na próxima página	

TÍTULO	CRITÉRIO DE SELEÇÃO
Personalized Online Education Learning Strategies Based on Transfer Learning Emotion Classification Model	Apresentar uma solução usando aprendizagem por transferência.
Improving the performance of automatic short answer grading using transfer learning and augmentation	Apresentar uma solução usando aprendizagem por transferência.

Fonte: Autoria Própria (2024).

B ARTIGOS INDISPONÍVEIS NA RSL

TÍTULO	DISPONIBILIDADE
Automatic question generation system for english reading comprehension	Indisponíveis.
A Deep Learning Model for an Intelligent Chat Bot System: An Application to E-Learning Domain	Indisponíveis.
Effects of schema acquisition and rule automation on mathematical problem-solving transfer.	Indisponíveis.
An Analysis of ResNet50 Model and RMSprop Optimizer for Education Platform Using an Intelligent Chatbot System	Indisponíveis.
Devoloping affordable tangible programming education applications using mobile vision	Indisponíveis.
Handwritten computer science words vocabulary recognition using concatenated convolutional neural networks	Indisponíveis.

Fonte: Autoria Própria (2024).

C DADOS EXTRAÍDOS DOS ARTIGOS

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Transfer Learning using Representation Learning in Massive Open Online Courses	Apresentar uma abordagem de aprendizado de transferência transdutiva automatizada que aborda o problema de desempenho de previsão ruim em MOOCs.	Aplicação de uma abordagem de aprendizado de transferência transdutiva automatizada baseada em representação de aprendizado com auto-encoders.	Aprendizado de transferência transdutiva automatizada baseada em representação de aprendizado com autoencoders, análise de componentes principais transdutiva, função de perda de alinhamento de correlação, uso de dados de cursos MOOCs e um problema de previsão de desistência e comparação com outras abordagens de aprendizado de transferência em MOOCs.	Cursos de ensino superior oferecidos pela internet.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Siamese-ERNIE: Topic Relevance Check for English Writing	Introduzir uma nova abordagem para avaliação de tópicos relevantes em escrita em inglês, utilizando aprendizado profundo e processamento de linguagem natural.	A metodologia utilizada no estudo é baseada em experimentos. Os autores coletaram um conjunto de dados e dividiram em amostras positivas e negativas e usaram o modelo Siamese-ERNIE para avaliar a relevância semântica entre o título e o conteúdo de cada ensaio e comparar o desempenho do modelo com outros modelos.	Os principais métodos e técnicas empregados no estudo são baseados em aprendizado profundo e processamento de linguagem natural.	O estudo não especifica o contexto educacional em que foi conduzido. Ele apenas menciona que o modelo foi treinado em um conjunto de dados com cerca de 2500 ensaios de estudantes do ensino médio.
Applying a Convolutional Neural Network to Screen for Specific Learning Disorder.	Avaliar a capacidade do aprendizado profundo usando transfer learning e o modelo MobileNetV2 para distinguir entre estudantes que têm um Transtorno de Aprendizagem Específico (SLD) E aqueles que não têm.	O estudo utiliza uma metodologia experimental. O estudo utiliza o modelo MobileNetV2 e transfer learning para treinar um algoritmo de aprendizado profundo para distinguir entre estudantes com e sem Transtorno de Aprendizagem Específico (SLD) Com base em amostras de escrita à mão.	Transfer learning, Deep learning, MobileNetV2, amostras de escrita à mão e medidas de desempenho.	O estudo foi conduzido no contexto educacional do ensino médio.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Unpacking the “Black Box” of AI in Education	Fornecer uma visão geral das técnicas de inteligência artificial (IA) atualmente utilizadas na educação, incluindo suas aplicações e limitações e propor soluções para os desafios e limitações atuais do uso de IA na educação, incluindo o uso de transfer learning e a necessidade de uma abordagem mais crítica e reflexiva ao desenvolvimento e implementação de sistemas de IA na educação.	O artigo não fornece informações específicas Sobre a metodologia usada no estudo.	O artigo discute várias técnicas de IA, incluindo aprendizado de máquina, redes neurais, algoritmos genéticos, sistemas especialistas, processamento de linguagem natural, entre outros. No entanto, o texto não se concentra em um método ou técnica específica, mas Sim em fornecer uma visão geral das técnicas de IA que podem ser usadas na educação.	O artigo discute como a IA pode ser usada em vários contextos educacionais, incluindo ensino fundamental, ensino superior e Treinamento corporativo.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
A Shallow System Prototype for Violent Action Detection in Italian Public Schools	Fornecer uma solução de baixo custo para detectar a violência nas escolas e melhorar a segurança dos alunos.	A metodologia utilizada no estudo é uma Combinação de experimento e análise quantitativa.	Rede neural convolucional 2D (CNN), Transfer learning, Conjunto de dados Daily School Break, Acurácia (A), Análise quantitativa e Análise qualitativa.	O estudo foi conduzido no contexto educacional de escolas públicas italianas, mas não especifica um nível de ensino específico, como Ensino fundamental, ensino superior ou treinamento corporativo.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
CNN feature and classifier fusion on novel transformed image dataset For dysgraphia diagnosis in children	Avaliar a eficácia da análise de imagens manuscritas, especificamente a análise de dados manuscritos offline, para diagnosticar a dislexia em crianças. Além disso, o estudo visa melhorar a precisão e eficácia dos modelos de diagnóstico de dislexia por meio de técnicas de aprendizado de máquina, como transfer learning e fusão de recursos. O estudo também compara o desempenho dos modelos propostos com os métodos de diagnóstico de dislexia Do estado da arte.	A metodologia usada no estudo inclui experimentos para avaliar a eficácia dos métodos propostos, bem como uma comparação de desempenho com Métodos de diagnóstico de disgrafias já existentes.	Transferência de aprendizado via ajuste fino e extração de recursos de redes neurais convolucionais (CNNs), fusão de recursos de diferentes CNNs, aprendizado em conjunto de múltiplos classificadores, técnicas de aumento de dados para aumentar o tamanho do conjunto de treinamento, utilização de classificadores de aprendizado de máquina tradicionais, como SVM, Random Forest e AdaBoost, bem como classificadores de aprendizado profundo, validação cruzada estratificada de dez vezes para avaliar o desempenho dos métodos propostos e comparação de desempenho com métodos de diagnóstico de disgrafia já existentes.	O estudo não especifica o contexto educacional em que foi conduzido.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
An Automated System for the Early Detection of Dysgraphia using Deep Learning Algorithms	Ajudar crianças com dificuldades de aprendizagem específicas, como a disortografia, e permitir intervenção e suporte oportunos.	O estudo apresentado neste artigo utiliza uma metodologia de análise quantitativa. Ele descreve a aplicação de métodos de aprendizado profundo para a detecção de disortografia em crianças, e apresenta resultados numéricos que demonstram a eficácia do sistema proposto.	Uso de algoritmos de aprendizado profundo, especificamente a técnica de transfer learning. Para a detecção de disortografia em crianças.	O artigo não especifica o contexto educacional em que o estudo foi conduzido. No entanto, o estudo é relevante para o contexto educacional em geral, pois a disortografia é uma dificuldade de aprendizagem que pode afetar crianças em diferentes níveis de ensino.
Deep Learning Applications for Dyslexia Prediction	Desenvolver modelos de aprendizagem automática precisos e confiáveis que possam auxiliar na detecção e diagnóstico precoce da dislexia, bem como obter uma melhor compreensão dos mecanismos neurais e cognitivos subjacentes à dislexia.	A metodologia utilizada no estudo foi uma revisão crítica da literatura sobre métodos de aprendizagem profunda para predição e classificação da dislexia.	Redes neurais artificiais (RNA), percepção multicamadas (MLP), máquina de vetores de suporte (SVM), eliminação recursiva de recursos com validação cruzada (RFE-CV), análise de componentes principais (PCA), estatísticas espaciais baseadas em trilhas (TBSS) e aprendizado profundo (DL) usando redes neurais convolucionais (CNN).	O estudo não se limita a um contexto educacional específico, mas visaram contribuir para um campo mais amplo de pesquisa e diagnóstico da dislexia.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Recognition of Student Engagement State in a Classroom Environment Using Deep and Efficient Transfer Learning Algorithm	Comparar o engajamento dos alunos em diferentes timestamps usando estatísticas inferenciais.	A metodologia usada neste estudo é uma combinação de experimento e análise quantitativa.	Coleta de dados, processamento de dados, modelo assistido por transfer learning, análise de dados, revisão da literatura e discussão dos resultados.	O estudo foi conduzido em um contexto de ensino superior, em um Ambiente de sala de aula offline.
A Survey on Recent Approaches to Question Difficulty Estimation from Text	Fornecer uma visão geral das abordagens existentes e avaliar sua eficácia em relação a diferentes tipos de perguntas e domínios educacionais.	A metodologia usada no estudo é uma revisão Sistemática da literatura.	Transfer Learning, extração de características e algoritmos de Aprendizado de Máquina.	O estudo não foi conduzido em um contexto educacional específico. Em vez disso, é uma revisão sistemática da literatura que analisa as abordagens existentes para a estimativa de dificuldade de perguntas em contextos educacionais usando técnicas de Processamento de Linguagem Natural (PLN).
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Improving the performance of automatic short answer grading using transfer learning and Augmentation	Discutir técnicas para melhorar o desempenho da avaliação automática de respostas curtas em um contexto educacional, Usando transfer learning e aumento de dados.	A metodologia usada neste estudo é baseada em experimentos comparativos. Vários experimentos comparativos foram conduzidos para avaliar a eficácia das estratégias de aumento de dados em tarefas de avaliação automática de respostas curtas em diferentes tamanhos de dados de treinamento.	Transfer learning e aumento de dados.	O artigo não especifica o contexto educacional exato em que o estudo foi conduzido. O artigo menciona que a avaliação automática de respostas curtas é uma tarefa importante em muitos contextos educacionais, incluindo ensino fundamental, ensino superior e treinamento corporativo.
Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods	Apresentar uma revisão da literatura sobre a automação da estimativa de engajamento em ambientes de aprendizagem inteligentes, incluindo definições de engajamento, conjuntos de dados e métodos baseados em aprendizado de máquina para a estimativa automática de engajamento.	A metodologia utilizada neste estudo é a revisão sistemática da literatura. Foram selecionados estudos de pesquisa publicados entre 2010 e 2022 que se concentraram em três questões de pesquisa relacionadas às definições de engajamento, conjuntos de dados e Métodos usados na literatura.	Revisão sistemática da literatura, a análise de definições de engajamento, a revisão de conjuntos de dados e a revisão de métodos baseados em aprendizado de máquina para a estimativa automática de engajamento. Além disso, o estudo apresenta uma análise quantitativa dos métodos de aprendizado de máquina mais utilizados na literatura, incluindo SVMs e técnicas de deep learning.	O artigo não fornece informações específicas sobre o contexto do ambiente educacional em que os estudos foram realizados.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Hand Gesture Recognition with ConvNets for School-Aged Children to Learn Basic Arithmetic Operations	Propor a construção de um Reconhecedor de Gestos de Mão capaz de identificar 13 classes.	Estudo de caso.	Pré-processamento de imagem, aprendizagem por transferência e treinamento de modelo.	Não é especificado o contexto educacional que foi conduzido, no entanto, o modelo central visa traduzir dígitos e símbolos matemáticos assinados por alunos e professores em um ambiente escolar para facilitar o processo de ensino-aprendizagem mais envolvente e promover a aprendizagem da língua de sinais na comunidade estudantil.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Bimodal Learning Engagement Recognition from Videos in the Classroom	Melhorar a eficiência do reconhecimento de comportamento dos estudantes e fornecer informações valiosas para os professores, a fim de melhorar a qualidade do ensino e Aprendizagem.	O estudo mencionado que utiliza uma metodologia de experimento.	Método de transfer learning profundo com VGG-16 para reconhecer o comportamento dos estudantes na sala de aula. Além disso, os autores utilizam técnicas de pré-processamento de imagem, como redimensionamento e normalização, para Preparar os dados de entrada para o modelo.	O estudo mencionado que não especifica o contexto educacional em que foi conduzido. No entanto, o estudo tem como objetivo reconhecer o comportamento dos estudantes na sala de aula, o que sugere que o contexto educacional pode ser de ensino fundamental, ensino médio Ou ensino superior.
Development of Low Resource Machine Learning Models for Child Cognitive Ability Assessments	Desenvolver soluções baseadas em modelos de aprendizado de máquina para a avaliação preliminar das habilidades cognitivas de crianças falantes de Sinhala, entre 7 e 9 anos De idade.	O estudo utiliza uma metodologia experimental, com o objetivo de desenvolver soluções baseadas em modelos de aprendizado de máquina para a avaliação preliminar das habilidades cognitivas de crianças falantes de Sinhala, entre 7 e 9 anos de idade.	Modelos de aprendizado de máquina e deep learning (DL), técnicas de reconhecimento de expressão facial e detecção de pontos de referência faciais e Transfer Learning para a classificação de formas desenhadas à mão.	O estudo não especifica um contexto educacional específico em que foi conduzido.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Deep Learning Based Emotion Recognition and Visualization of Figural Representation	Propor um algoritmo de rede neural aprimorado - CNN-BiLSTM para melhorar o desempenho do reconhecimento de emoções e conduzir um experimento de simulação para verificar o desempenho do algoritmo proposto.	A metodologia utilizada no estudo é a análise Experimental e a pesquisa.	Algoritmo CNN-BiLSTM, comparação de desempenho da abordagem proposta com outros modelos de aprendizado de máquina, como SVM, SVM-RBF, ELM e CNN, usando diversas métricas, como precisão e taxas de Recall.	O estudo não especifica o contexto educacional exato em que o experimento foi conduzido.
Meta Transfer Learning for Early Success Prediction In MOOCs	Os principais objetivos do estudo são explorar o uso de transfer learning para prever o sucesso dos alunos em MOOCs, apresentar e analisar três estratégias para criar modelos generalizáveis para essa tarefa e avaliar o desempenho desses modelos em diferentes cenários.	A metodologia usada no estudo é baseada em experimentos quantitativos.	Transfer learning, aprendizado de máquina, pré-processamento de dados, e realizaram análises estatísticas para avaliar a eficácia das estratégias propostas.	O estudo foi conduzido no contexto de MOOCs (Massive Open Online Courses), que são cursos online abertos e massivos, conduzido no contexto do ensino superior.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Classification in Math Class: Using Convolutional Neural Networks to Categorize Student Cognitive Demand	Explorar como a visão computacional pode ajudar os professores a medir e acompanhar a demanda cognitiva dos alunos durante tarefas rigorosas de matemática, bem como investigar como o uso de técnicas de transfer learning pode melhorar a precisão da classificação.	Análise quantitativa e qualitativa para investigar a qualidade da instrução em matemática.	Análise de gravações de vídeo em sala de aula de matemática de oitava série, o uso de técnicas de visão computacional e transfer learning para categorizar a demanda cognitiva dos alunos e a utilização da ferramenta Instructional Quality of Assessment Rubric para avaliar a qualidade da instrução em Matemática.	O estudo foi conduzido em salas de aula de matemática de oitava série, o que sugere que o contexto educacional é o Ensino Fundamental II.
Generating Answer Candidates for Quizzes and Answer-Aware Question Generators	Propor um modelo que possa gerar candidatos a respostas para perguntas de quiz abertas, tornando a tarefa de preparação de perguntas mais fácil e eficiente.	A metodologia usada neste estudo é experimental, com a realização de experimentos para avaliar o desempenho do modelo proposto em comparação com outros métodos de geração de perguntas e Respostas.	O artigo menciona o uso de modelos pré-treinados, como BERT, RoBERTa, ALBERT, T5 e BART, para gerar perguntas e avaliar os resultados usando métricas como BLEU, ROUGE e METEOR.	O estudo não especifica um contexto educacional específico em que foi conduzido. Acredita-se que a abordagem proposta possa ser aplicada em vários contextos educacionais, incluindo ensino Fundamental, ensino superior e treinamento corporativo.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
A deep transfer learning approach to modeling teacher discourse in the Classroom	Explorar a transferência de aprendizado para adaptar modelos treinados em um conjunto de dados para outra tarefa relacionada, como a análise do discurso do professor.	A metodologia usada neste estudo é principalmente experimental e envolve a comparação de duas abordagens de aprendizado de máquina para modelar sete características do discurso do professor. Os autores também realizam uma análise quantitativa dos resultados, comparando a precisão das duas abordagens em diferentes quantidades de dados de treinamento e em diferentes níveis de granularidade de Feedback.	Abordagens de aprendizado de máquina, transferência de aprendizado, análise Quantitativa e revisão da literatura.	O estudo não especifica o contexto educacional em que foi conduzido.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Modeling review helpfulness with augmented transfer neural Networks	Investigar a detecção automática de características de feedback úteis em comentários de revisão de alunos.	A metodologia utilizada no estudo é principalmente a análise quantitativa de um conjunto de dados de comentários de feedback de uma plataforma educacional online.	Aprendizagem por transferência com BERT e um modelo de rede neural de transformador aumentado, para analisar as características dos comentários de feedback e prever sua utilidade. O estudo também utiliza análise estatística para identificar as características dos comentários de feedback que estão associadas às percepções dos alunos.	O estudo foi conduzido no contexto de uma aula de graduação/pós-graduação em ciência da computação de nível superior, baseada em projetos e revisada por pares.
Personalized Online Education Learning Strategies Based on Transfer Learning Emotion Classification Model	Propor um modelo de rede sensível ao contexto baseado em transfer learning para prever o desempenho dos alunos em cursos online e desenvolver estratégias de aprendizagem personalizadas.	A metodologia usada no estudo é um estudo de caso experimental. O estudo foi conduzido em duas turmas de cursos MOOC e os dados foram pré-processados, incluindo extração de recursos de dados de Clickstream de vídeo.	Transfer learning, extração de características baseadas em seleção livre, visualização de interações de vídeo MOOC e desenvolvimento de modelos de classificação afetiva para implementar estratégias de aprendizagem online personalizadas.	O estudo foi conduzido no contexto de cursos MOOC (Massive Open Online Courses), que são cursos online abertos e massivos. Não há menção específica sobre o nível educacional dos alunos que participaram dos cursos MOOC.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Math Word Problem Solver Based On Text-to-Text Transformer Model	Propor um modelo de aprendizado profundo baseado em transformador de texto para resolver problemas matemáticos em linguagem natural e avaliar a eficácia do modelo proposto em comparação com outros métodos de resolução de problemas Matemáticos.	A metodologia usada no estudo é principalmente experimental, com o objetivo de avaliar a eficácia do modelo proposto em resolver problemas matemáticos em linguagem natural.	Aprendizado profundo, pré-treinamento em um corpus chinês usando uma técnica de modelagem de linguagem natural, que é uma técnica comum em transfer learning, classificação dos problemas matemáticos, experimentos para avaliar a eficácia do modelo proposto em comparação com outros métodos de resolução de problemas matemáticos, análise quantitativa dos resultados experimentais e análise qualitativa para explicar as diferenças observadas na eficácia do modelo em diferentes categorias de Problemas matemáticos.	O estudo não especifica um contexto educacional específico em que foi conduzido.
Development of CNN Transfer Learning for Dyslexia Handwriting Recognition	Contribuir para a melhoria da detecção de dislexia e apoio aos alunos disléxicos.	O estudo descrito no artigo utiliza uma metodologia experimental.	Transfer Learning, convolutional Neural Network (CNN), LeNet-5 e Jupyter Notebook.	O contexto educacional em que o estudo foi conduzido não é mencionado no artigo.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
DRAFT: A Novel Framework to Reduce Domain Shifting in Self-supervised Learning and Its Application to Children's ASR	Introduzir métodos de aprendizado auto-supervisionado (SSL) para tarefas de reconhecimento automático de fala (ASR) para crianças, utilizando dados de fala de adultos não anotados na fase de pré-treinamento.	A metodologia usada neste estudo é principalmente experimental.	Aprendizado auto-supervisionado (SSL), uma nova estrutura de adaptação responsável pelo domínio (DRAFT), Speed perturbation e SpecAug, Avaliação da eficácia da estrutura DRAFT em tarefas de ASR para crianças, experimentos de transferência de conhecimento cruzado da estrutura DRAFT para outras tarefas de ASR de baixo recurso e busca de decodificação por busca gulosa durante a avaliação.	O estudo não foi conduzido em um contexto educacional específico, mas sim em um contexto de pesquisa em reconhecimento automático de fala (ASR) para crianças.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
EXAMS: A Multi-Subject High School Examinations Dataset for Cross-Lingual and Multilingual Question Answering	O estudo apresenta um conjunto de dados de referência para perguntas e respostas multilíngues e interlinguais para exames de ensino médio, com mais de 24.000 perguntas de exame em 16 idiomas, cobrindo 8 famílias de idiomas e 24 disciplinas escolares.	A metodologia usada no artigo não é claramente especificada.	Foram realizados experimentos com modelos pré-treinados multilíngues existentes para avaliar o desempenho em relação ao conjunto de dados Exams.	O contexto educacional em que o estudo foi conduzido é o ensino médio.
Spontaneous facial expression database for academic emotion Inference in online learning	Propor um método de transfer learning para treinar um modelo de rede neural convolucional para inferir emoções acadêmicas a partir de expressões faciais.	A metodologia usada no estudo é principalmente experimental.	Coleta de dados, anotação de dados, transfer learning, data augmentation, análise de Correlação e avaliação de desempenho.	O estudo foi conduzido em um contexto educacional de Ensino Superior, com participantes universitários.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Relationships among cognitive components in Logo learning and Transfer.	Aprofundar a compreensão dos processos cognitivos envolvidos na aprendizagem de programação em Logo e como esses processos podem ser transferidos para outras tarefas e contextos.	A metodologia usada no estudo foi uma combinação de experimento e análise quantitativa.	Instrução mediada, avaliação do componente cognitivo e avaliação da tarefa de transferência.	O estudo não menciona explicitamente o contexto educacional em que foi realizado.
Learning and Transfer Effects of Embodied Simulations Targeting Cross-cutting Concepts in Science	Explorar como atividades físicas podem melhorar a compreensão dos alunos sobre conceitos científicos complexos e como esses conceitos podem ser transferidos para novos tópicos.	A metodologia usada neste estudo é baseada em experimentos.	Experimentos para avaliar como atividades físicas, medidas e métodos analíticos para avaliar o engajamento dos participantes e os resultados dos experimentos e revisão da literatura para discutir a importância da aprendizagem corporificada.	O estudo foi conduzido em um contexto educacional de ensino superior.
Continua na próxima página				

TÍTULO	OBJETIVOS	METODOLOGIA	MÉTODOS E TÉCNICAS	CONTEXTO EDUCACIONAL
Interdomain Transfer Between Isomorphic Topics in Algebra And Physics	Investigar a transferência de resolução de problemas com base na capacidade de aplicar procedimentos matemáticos a problemas simples de palavras em álgebra e física.	A metodologia usada no estudo é um experimento.	Seleção aleatória de participantes, testes pré e pós-estudo, estudo de capítulo, transferência de problemas, análise quantitativa e categorização de soluções.	O estudo foi conduzido em um contexto educacional de ensino médio.
Depth and Breadth of Vocabulary in Two Languages: Which Vocabulary Skills Transfer?	Examinar a relação entre o conhecimento de palavras paradigmáticas e sintagmáticas em espanhol e inglês em alunos bilíngues da 4ª e 5ª séries, e explorar a transferência de habilidades de vocabulário entre as duas línguas.	A metodologia usada neste estudo foi uma análise quantitativa dos dados coletados de um teste de definição e descrição de palavras semelhantes em espanhol e inglês, administrado a 88 alunos bilíngues da 4ª e 5ª séries.	Teste de vocabulário, teste de compreensão de leitura, tarefa de definição e descrição de palavras e análise quantitativa.	O estudo foi conduzido em escolas primárias públicas nos Estados Unidos, com alunos bilíngues da 4ª e 5ª séries. Portanto, o contexto educacional em que o estudo foi conduzido é o ensino fundamental.

Fonte: Autoria Própria (2024).

D DATASETS UTILIZADOS NOS ARTIGOS

TÍTULO	DATASETS UTILIZADOS	FONTE
Transfer Learning using Representation Learning in Massive Open Online Courses	Os dados utilizados na pesquisa estão relacionados aos Massive Open Online Courses (MOOCs) e incluem dados de séries temporais de tipos específicos de atividades MOOC por aluno.	-
Siamese-ERNIE: Topic Relevance Check for English Writing	Cerca de 2500 ensaios de estudantes do Ensino Médio, coletados aleatoriamente online, abrangendo diferentes tópicos.	-
Applying a Convolutional Neural Network to Screen for Specific Learning Disorder.	O artigo não menciona especificamente quais datasets foram utilizados no estudo. No entanto, o texto menciona que amostras de escrita à mão foram coletadas dos estudantes e usadas como entrada para o modelo de aprendizado profundo.	-
Unpacking the “Black Box” of AI in Education	-	-
A Shallow System Prototype for Violent Action Detection in Italian Public Schools	O estudo utilizou o conjunto de dados Daily School Break, que foi criado especificamente para o estudo e contém vídeos originais gravados no pátio de uma Escola italiana.	O conjunto de dados Daily School Break foi criado especificamente para este estudo e não foi obtido de nenhuma fonte externa. Os próprios autores coletaram os vídeos gravando as atividades no pátio de uma escola pública na Itália durante a manhã, usando diferentes ângulos amplos.
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children	O estudo utilizou um conjunto de dados online para diagnóstico de disgrafia em crianças, que consiste em dados manuscritos adquiridos para seis atividades de escrita diferentes em ortografia eslovaca. O conjunto de dados contém amostras de escrita de 120 alunos, dos quais 63 apresentam escrita normalmente desenvolvida e os 57 restantes apresentam disgrafia.	O conjunto de dados utilizado no estudo foi obtido a partir do único conjunto de dados online disponível publicamente para o problema de diagnóstico de disgrafia. Ele foi criado pelos autores Drotár e Dobeš e consiste em dados manuscritos adquiridos para seis atividades de escrita diferentes em ortografia eslovaca.
An Automated System for the Early Detection of Dysgraphia using Deep Learning Algorithms	Um conjunto de dados foi usado para a escrita manual padrão e outro para a escrita ilegível. Para a escrita manual padrão, os dados foram extraídos do banco de dados NIST Special Database 19, que inclui 3600 formulários de amostra manuscritos de escritores, 810.000 imagens de caracteres isolados desses formulários, classificações de verdade básica para os dados de imagem e formulários de referência para coleta de dados adicionais. Para a escrita ilegível, foram utilizadas imagens de escrita manual espelhadas. O conjunto de dados final incluiu cerca de 2000-3000 imagens para cada letra, com cada imagem medindo 28x28 pixels.	O banco de dados utilizado neste estudo é o NIST Special Database 19, que é um conjunto de dados público disponibilizado pelo National Institute of Standards and Technology (NIST) dos Estados Unidos.
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Deep Learning Applications for Dyslexia Prediction	Conjuntos de dados de escrita manual: amostras de escrita manual de crianças disléxicas e não disléxicas. Conjuntos de dados de ressonância magnética e fMRI: imagens cerebrais de crianças disléxicas e não disléxicas. Conjuntos de dados de avaliação comportamental e cognitiva: informações sobre o desempenho de crianças disléxicas e não disléxicas em diversas Tarefas cognitivas.	-
Recognition of Student Engagement State in a Classroom Environment Using Deep and Efficient Transfer Learning Algorithm	Dois datasets foram utilizados neste estudo: um dataset coletado durante a palestra em um ambiente de sala de aula offline e outro dataset já disponível, como o DAiSEE e o EmotiW, que foram usados para treinar e testar outras ferramentas de detecção de engajamento de alunos.	-
A Survey on Recent Approaches to Question Difficulty Estimation from Text	Conjuntos de dados de propósito geral, como a Wikipedia, enquanto os modelos construídos para CKA (avaliação de conhecimento) aproveitam conjuntos de dados relacionados aos tópicos avaliados pelas perguntas, como livros e transcrições de palestras. Conjuntos de dados experimentais usados para avaliar o desempenho do modelo varia amplamente nos diferentes trabalhos revisados.	-
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Improving the performance of automatic short answer grading using transfer learning and Augmentation	O estudo utilizou um conjunto de dados chamado Short Programming Related Answer Grading Dataset (SPRAG), que contém respostas de referência e de estudantes envolvendo conceitos e bibliotecas de programação Python. O conjunto de dados contém 4039 registros (2707 corretos e 1332 incorretos) e respostas a 114 perguntas diferentes, distribuídas em cinco categorias diferentes.	-
Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods	Conjunto de dados de seminário em sala de aula IIITB, conjunto de dados de seminário on-line IIITB, conjunto de dados de estilo de apresentação IIITB, LectureVideoDB, ClassX, IIIT-AR-13K e muito mais.	-
Hand Gesture Recognition with ConvNets for School-Aged Children to Learn Basic Arithmetic Operations	Conjunto de dados COCO 2017.	O artigo menciona que esse conjunto de dados pertence ao TensorFlow 2 Detection Model Zoo.
Bimodal Learning Engagement Recognition from Videos in the Classroom	-	-
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Development of Low Resource Machine Learning Models for Child Cognitive Ability Assessments	O estudo utilizou três conjuntos de dados: o conjunto de dados de letras cingalesas, que continha 179.886 imagens de 28x28 pixels de letras cingalesas manuscritas e geradas por computador pertencentes a 16 classes; o conjunto de dados geométricos de formas 2D, que continha 16.208 imagens de 224x224 pixels de 4 classes de formas geométricas; e o conjunto de dados de formas desenhadas à mão, que continha 3.000 imagens de 224x224 pixels de 3 classes de formas desenhadas à mão. Para reconhecimento de expressões faciais, o estudo utilizou o conjunto de dados de expressões Faciais de crianças LIRIS-CSE.	O banco de dados Sinhala letter dataset utilizado no estudo foi uma coleção de letras Sinhala desenhadas à mão por crianças em idade escolar primária do Sri Lanka e letras geradas por computador devido à quantidade insuficiente de dados coletados das crianças. As letras geradas por computador foram criadas usando uma biblioteca gratuita e de código aberto chamada Python Imaging Library, que usou sete tipos de fontes que se relacionam principalmente com letras desenhadas à mão. O conjunto de dados de expressão facial LIRIS-CSE children foi coletado pelo Laboratório de Informática de Imagens e Sinais de Grenoble (LIRIS) e pelo Centro de Sistemas Embarcados (CSE) da École Centrale de Lyon.
Deep Learning Based Emotion Recognition and Visualization of Figural Representation	O estudo não menciona nenhum conjunto de dados específico usado para treinar ou testar o algoritmo CNN-BiLSTM. No entanto, os autores citam um levantamento de conjuntos de dados, recursos e algoritmos para análise de microexpressão facial baseada em vídeo, que pode ser relevante para a abordagem do estudo. A pesquisa discute vários conjuntos de dados, como SMIC, CASME e SAMM, que são comumente usados para análise de Expressões faciais.	-
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Meta Transfer Learning for Early Success Prediction In MOOCs	O estudo utilizou um dataset que incluiu 26 MOOCs ministrados por uma universidade europeia entre 2013 e 2015. O dataset foi totalmente anonimizado em relação às informações dos alunos e contém interações de 145.714 alunos. Após a remoção de alunos que abandonaram o curso precocemente, o Dataset final contém 73.042 alunos.	O banco de dados utilizado no estudo foi coletado de 26 MOOCs ministrados por uma universidade europeia entre 2013 e 2015.
Classification in Math Class: Using Convolutional Neural Networks to Categorize Student Cognitive Demand	O estudo utilizou dados coletados de quatro gravações de vídeo de 30 minutos cada, realizadas em salas de aula de matemática de oitava série, durante a primavera de 2017.	-
Generating Answer Candidates for Quizzes and Answer-Aware Question Generators	O estudo menciona vários datasets que foram utilizados em pesquisas relacionadas à geração automática de perguntas e respostas, incluindo Crowdsourcing, ARC, Open-BookQA, exames de múltipla escolha em búlgaro e vietnamita e EXAMS, que cobre 16 idiomas diferentes. No entanto, o estudo não especifica quais desses datasets foram usados especificamente para avaliar a abordagem Proposta neste artigo.	-
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
A deep transfer learning approach to modeling teacher discourse in the Classroom	O estudo utiliza um conjunto de dados específico do domínio, composto por 16.977 amostras de áudio e transcrições de aulas de ciências do ensino médio. As amostras foram coletadas de 94 aulas ministradas por 12 professores diferentes em escolas públicas dos Estados Unidos.	O conjunto de dados utilizado neste estudo foi coletado pelos próprios autores. Eles recrutaram 16 professores de Língua Inglesa (ELA) de três distritos escolares suburbanos da Pensilvânia e os treinaram para gravar de forma independente suas próprias palestras em sala de aula. Cada professor gravou pelo menos quatro sessões de duas turmas diferentes, resultando num total de 142 gravações originais. Destes, os autores identificaram um total de 127 gravações que podiam ser utilizadas para análise automatizada. As gravações foram segmentadas e transcritas automaticamente por meio do reconhecedor de fala IBM Watson, e um subconjunto de 16.977 enunciados codificados foi utilizado para análise.
Modeling review helpfulness with augmented transformer neural Networks	O estudo usa um conjunto de dados de comentários de uma plataforma educacional online chamada Expertiza. O conjunto de dados contém 50.000 comentários de 1.000 alunos em 10 turmas diferentes.	O conjunto de dados utilizado no estudo vem de uma plataforma educacional online chamada Expertiza. Expertiza é uma plataforma de código aberto desenvolvida no Georgia Institute of Technology para apoiar a aprendizagem baseada em projetos e revisão por pares. A plataforma permite que os alunos enviem seus trabalhos e recebam feedback de seus colegas e instrutores.
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Personalized Online Education Learning Strategies Based on Transfer Learning Emotion Classification Model	-	-
Math Word Problem Solver Based On Text-to-Text Transformer Model	O estudo utilizou um conjunto de dados de 210.000 problemas matemáticos em chinês, que foram coletados de livros didáticos de escolas primárias na China.	O banco de dados utilizado no estudo foi coletado de livros didáticos de escolas primárias na China.
Development of CNN Transfer Learning for Dyslexia Handwriting Recognition	O artigo menciona que foram utilizados dois datasets: um dataset de treinamento e um dataset de teste. No entanto, não são especificados os nomes ou fontes desses datasets.	-
DRAFT: A Novel Framework to Reduce Domain Shifting in Self-supervised Learning and Its Application to Children's ASR	OGI - Contém fala de crianças do jardim de infância até a 10ª série. O conjunto de dados é dividido em conjuntos de treinamento (70%), desenvolvimento (15%) e teste (15%) sem sobreposição de alto-falante. Cerca de 50 horas de dados de fala de crianças são usados para treinar o sistema ASR para crianças. MyST - Contém fala de conversação entre crianças e um tutor virtual de alunos da 3ª à 5ª série. O conjunto de dados contém 499 horas de dados de fala com 244.069 enunciados. No entanto, apenas 42% do corpus (240 horas) é anotado para ASR. O conjunto de dados também contém conjuntos de Desenvolvimento e teste para avaliação.	OGI Kids' Speech Corpus (OGI) - Este banco de dados foi coletado pela Universidade de Oregon e contém fala de crianças do jardim de infância até a 10ª série. My Science Tutor (MyST) - Este banco de dados foi coletado pela Universidade de Memphis e contém fala de conversação entre crianças e um tutor virtual de alunos da 3ª à 5ª série. O conjunto de dados contém 499 horas de dados de fala com 244.069 enunciados. No entanto, apenas 42% do corpus (240 horas) é anotado para ASR.
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
EXAMS: A Multi-Subject High School Examinations Dataset for Cross-Lingual and Multilingual Question Answering	O estudo apresenta um novo conjunto de dados e referência para perguntas e respostas multilíngues e interlinguais para exames de ensino médio, chamado de Exams. Além disso, é mencionado outros conjuntos de dados, como SQuAD, MLQA, XQuAD e TyDi QA, que foram usados para avaliar o desempenho de modelos pré-treinados multilíngues. Existem em relação ao conjunto de dados Exams.	O conjunto de dados Exams foi coletado a partir de exames escolares oficiais preparados pelos ministérios da educação de vários países.
Spontaneous facial expression database for academic emotion Inference in online learning	O estudo propôs a criação de um novo conjunto de dados chamado OL-SFED (Online Learning Spontaneous Facial Expression Database), que inclui vídeos e imagens faciais estáticas de expressões faciais espontâneas de participantes universitários enquanto assistiam a vídeos de cursos online. O OL-SFED é um conjunto de dados inédito e foi criado especificamente para o estudo.	O banco de dados, denominado OL-SFED (Online Learning Spontaneous Facial Expression Database), foi criado especificamente para este estudo. Os participantes eram estudantes universitários que foram gravados enquanto assistiam aos vídeos do curso online. Todos os participantes do banco de dados assinaram o termo de consentimento concordando com o uso de suas imagens faciais para fins de pesquisa. Os autores pretendem disponibilizar publicamente o OL-SFED e o acesso à base de dados pode ser solicitado por e-mail aos autores correspondentes.
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Relationships among cognitive components in Logo learning and Transfer.	O estudo coletou dados através de várias medidas, incluindo avaliações de componentes cognitivos e tarefas de transferência, bem como uma tarefa de classificação de cartões. Os dados foram então analisados por meio de modelagem de equações estruturais e análise de caminhos para identificar relações entre as variáveis.	-
Learning and Transfer Effects of Embodied Simulations Targeting Crosscutting Concepts in Science	O estudo utilizou dados coletados de dois experimentos separados, cada um com seu próprio conjunto de dados. O experimento 1 coletou dados de 51 estudantes do ensino médio, dos quais 34 tiveram seus dados incluídos nas análises. O experimento 2 coletou dados de 49 estudantes universitários, dos quais 46 tiveram seus dados incluídos nas análises.	-
Interdomain Transfer Between Isomorphic Topics in Algebra And Physics	Os pesquisadores criaram seis conjuntos de problemas de palavras paralelos para uso em sessões de pré-teste, pós-teste e transferência. Cada conjunto continha dois problemas de sequência aritmética e dois problemas de série aritmética, todos equilibrados em relação às variáveis fornecidas e às variáveis que precisavam ser resolvidas. Os conjuntos foram construídos especificamente para o estudo e não foram retirados de nenhum dataset existente.	-
Continua na próxima página		

TÍTULO	DATASETS UTILIZADOS	FONTE
Depth and Breadth of Vocabulary in Two Languages: Which Vocabulary Skills Transfer?	O estudo utilizou apenas um conjunto de dados, que foi coletado como parte de um estudo maior sobre o desenvolvimento de vocabulário entre alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos. O conjunto de dados incluiu informações sobre 88 alunos, 38 da 4ª série e 50 da 5ª série, que foram submetidos a testes de vocabulário em inglês e espanhol.	O estudo examinou a transferência de habilidades de vocabulário entre duas línguas em alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos.

Fonte: Autoria Própria (2024).

E DADOS (OUTROS) EXTRAÍDOS DOS ARTIGOS

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Transfer Learning using Representation Learning in Massive Open Online Courses	Método de aprendizagem por transferência baseado em representação que usa uma representação de espaço latente aprendida automaticamente usando dados de origem e destino.	Os autores usam uma representação de espaço latente aprendida automaticamente usando dados de origem e de destino. Eles avaliam a transferibilidade desta abordagem nas ofertas de dois cursos e entre dois cursos. Os autores usam a previsão de evasão como exemplo de um problema de modelagem preditiva e comparam e contrastam os cursos usados para avaliar os métodos de aprendizagem por transferência.	Previsão de evasão e comparar e contrastar os cursos usados para avaliar os métodos de aprendizagem por transferência.	As pontuações da AUC de previsão de abandono melhoram em 8% usando qualquer um dos métodos em comparação com a linha de base da transferência ingênua. Nesse sentido, os autores acreditam ter feito progressos significativos na resolução do problema de transferência de aprendizagem em MOOCs.	Os autores usam a área sob a curva característica de operação do receptor (AUC) para medir o desempenho de todos os modelos preditivos.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Siamese-ERNIE: Topic Relevance Check for English Writing	O estudo utiliza uma abordagem de transferência de aprendizado (transfer learning) com o modelo Siamese-ERNIE, que é uma rede neural pré-treinada baseada no modelo ERNIE (Enhanced Representation through Knowledge Integration), desenvolvido pela Baidu.	No estudo, o transfer learning foi aplicado por meio do modelo Siamese-ERNIE, que é uma rede neural pré-treinada baseada no modelo ERNIE. O Siamese-ERNIE é treinado em um conjunto de dados específico e, em seguida, é ajustado (fine-tuned) para a tarefa de classificação de similaridade de texto.	Uso do modelo Siamese-ERNIE, que é uma rede neural pré-treinada baseada no modelo ERNIE.	O estudo mostrou que o modelo alcançou uma precisão de 62% e um NDCG@4 de 0,72 na tarefa de similaridade. O estudo também comparou o desempenho do modelo com diferentes tamanhos de camada MLP e mostrou que a precisão e o NDCG diminuem à medida que o tamanho da camada MLP diminui.	A acurácia e o NDCG (Normalized Discounted Cumulative Gain). O estudo comparou o desempenho do modelo com diferentes tamanhos de camada MLP (Multi-Layer Perceptron) e mostrou que a precisão e o NDCG diminuem à medida que o tamanho da camada MLP diminui.
Applying a Convolutional Neural Network to Screen for Specific Learning Disorder.	Uma Rede Neural Convolutiva (CNN).	O transfer learning foi aplicado neste estudo utilizando o modelo MobileNetV2 pré-treinado como ponto de partida. A última camada do modelo foi removida e novas camadas foram adicionadas para se adequar à tarefa específica de detecção de SLD com base em amostras de escrita à mão.	Aprendizagem por transferência com um modelo MobileNetV2 pré-treinado e, em seguida, ajustá-lo removendo a última camada e adicionando novas camadas adequadas para a tarefa específica de detecção de SLD com base na caligrafia. amostras. O modelo foi então treinado em um conjunto de dados de amostras de caligrafia coletadas de estudantes do ensino médio.	AUC = 0,89, precisão = 0,94, recall = 0,89, score F = 0,91 e precisão = 0,92. Esses resultados indicam que o modelo possui um alto nível de precisão na detecção de SLD com base em amostras de caligrafia.	Área sob a curva (AUC), precisão, recall, pontuação F e exatidão. Essas métricas foram calculadas com base nos resultados obtidos no conjunto de validação, que consistiu em 50 imagens de caligrafia.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Unpacking the “Black Box” of AI in Education	O artigo menciona vários avanços algorítmicos, incluindo transferência de aprendizagem. O artigo discute várias arquiteturas de IA, incluindo RNNs, Transformers e CNNs, e como elas processam e transformam entradas em saídas.	O artigo menciona a aprendizagem por transferência como um avanço algorítmico que pode ajudar nos casos em que grandes conjuntos de dados não estão disponíveis para uma tarefa específica. No entanto, o artigo não fornece nenhum estudo específico ou aplicação da aprendizagem por transferência na educação.	Capacidades preditivas de modelos pré-treinados em dados de ambientes escolares maiores como ponto de partida.	O artigo não fornece nenhum estudo específico ou aplicação de IA na educação, nem menciona quaisquer medidas utilizadas para avaliar o desempenho de um modelo de IA na educação.	O artigo discute a importância de avaliar a eficácia e o impacto dos sistemas de IA na educação e destaca a necessidade de transparência e responsabilidade no desenvolvimento e implantação destes sistemas.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
A Shallow System Prototype for Violent Action Detection in Italian Public Schools	Os autores utilizaram técnicas de aprendizagem por transferência e redes neurais pré-treinadas, especificamente o modelo VGG19.	O processo de transfer learning permitiu que o modelo analisasse uma nova imagem, aproveitando as propriedades do conhecimento adquirido anteriormente para reconhecer novos objetos ou ações.	O processo de transfer learning foi realizado através da reutilização de pesos de uma rede neural pré-treinada, a VGG16, que foi treinada em um grande conjunto de dados de imagens. Em seguida, a rede neural foi treinada com o conjunto de dados Daily School Break, que contém vídeos originais gravados em um pátio de escola italiana.	Os resultados experimentais mostram que ao refinar o conjunto de dados via web scraping e aplicar técnicas de transferência de aprendizagem, a classificação da rede atinge uma precisão em torno de 95%. No entanto, o artigo também menciona que alguns obstáculos ainda permanecem, como o ruído nas imagens e a dificuldade de obtenção de imagens tendo menores como sujeitos principais.	Usaram precisão, recall e pontuação F1 para comparar os resultados obtidos em arquiteturas de baixo custo.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
CNN feature and classifier fusion on novel transformed image dataset for dysgraphia diagnosis in children	Transfer learning via fine-tuning que foram desenvolvidas redes neurais convolucionais (CNNs) específicas para cada tarefa. Extração de características de CNN que foram extraídas do conjunto de dados específico da tarefa usando uma rede DenseNet201 pré-treinada.	Os autores utilizaram a popular arquitetura de rede neural chamada DenseNet201, que foi pré-treinada no conjunto de dados ImageNet, para implementar a aprendizagem por transferência. O modelo DenseNet201 pré-treinado foi ajustado em dados de imagens manuscritas específicas de tarefas para desenvolver redes neurais convolucionais (CNNs) específicas de tarefas. Esses modelos DenseNet ajustados foram então utilizados como classificadores básicos para aprendizagem em conjunto. Além disso, o modelo DenseNet201 pré-treinado foi usado como um extrator de recursos para extrair recursos CNN dos dados de imagem manuscrita específicos da tarefa.	Utilizaram a arquitetura de rede neural chamada DenseNet201, que foi pré-treinada no conjunto de dados ImageNet, para implementar a aprendizagem por transferência. O modelo DenseNet201 pré-treinado foi ajustado em dados de imagens manuscritas específicas de tarefas para desenvolver redes neurais convolucionais (CNNs) específicas de tarefas. Esses modelos DenseNet ajustados foram então utilizados como classificadores básicos para aprendizagem em conjunto. Além disso, o modelo DenseNet201 pré-treinado foi usado como um extrator de recursos para extrair recursos CNN dos dados de imagem manuscrita específicos da tarefa.	A abordagem proposta melhorou significativamente o desempenho da classificação, alcançando uma precisão de 97,3%. Esta melhoria substancial demonstra a eficácia dos métodos propostos. Além disso, a adoção de técnicas de fusão de recursos e aprendizagem de conjunto desempenhou um papel significativo no desenvolvimento de um algoritmo inteligente de tomada de decisão com uma precisão de 97%.	Para avaliar o desempenho, os autores realizaram vários experimentos. Um total de 45 classificadores de ML tradicionais (incluindo SVM, Random Forest e AdaBoost treinados em 15 conjuntos de recursos diferentes) e 20 classificadores de aprendizagem profunda (compreendendo quatro CNNs profundos ajustados e 15 combinações possíveis de conjuntos de CNNs profundos ajustados) foram treinados e avaliado por meio de validação cruzada estratificada de dez vezes. A avaliação ou avaliação dos métodos propostos é crucial para examinar a sua eficácia na abordagem do problema em questão. Para implementar algoritmos tradicionais de aprendizado de máquina, foi utilizada a popular estrutura de aprendizado de máquina SciKit.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
An Automated System for the Early Detection of Dysgraphia using Deep Learning Algorithms	A solução apresentada neste artigo envolve o uso de algoritmos de aprendizagem profunda, especificamente redes neurais convolucionais (CNNs) e aprendizagem por transferência.	Para aplicar a aprendizagem por transferência foram utilizados modelos de redes neurais convolucionais pré-treinados. Esses modelos foram treinados em grandes conjuntos de dados para tarefas de classificação de imagens, e seus recursos aprendidos foram usados como ponto de partida para a tarefa de detecção de disgrafia. Os modelos pré-treinados foram ajustados no conjunto de dados de disgrafia para se adaptarem às características específicas da caligrafia disgráfica. Ao reutilizar os recursos aprendidos nos modelos pré-treinados, o sistema conseguiu obter maior precisão na detecção de disgrafia com menos amostras de treinamento.	Foram utilizados modelos de redes neurais convolucionais pré-treinados, como VGG16, Inception V3 e Resnet. Esses modelos foram ajustados no conjunto de dados de disgrafia para se adaptarem às características específicas da caligrafia disgráfica. Ao reutilizar os recursos aprendidos nos modelos pré-treinados, o sistema conseguiu obter maior precisão na detecção de disgrafia com menos amostras de treinamento. O artigo menciona que as camadas inferiores dos modelos pré-treinados são mais específicas da aplicação, enquanto as camadas superiores adquirem intrinsecamente propriedades básicas como bordas e texturas.	A precisão geral obtida para o Inception V3 foi em torno de 97,94%, para o VGG16 foi em torno de 97,98% e para o Resnet foi em torno de 98,22%. Estes resultados sugerem que os modelos de aprendizagem profunda são eficazes na detecção de disgrafia com alta precisão. O artigo apresenta que mostra o desempenho do modelo CNN, que alcançou precisão de 85% para detecção de disgrafia.	Para avaliar o desempenho dos modelos de aprendizagem por transferência, o artigo relata a precisão geral obtida para cada modelo. O artigo menciona que também foram calculadas a sensibilidade e a especificidade dos modelos. O artigo também apresenta uma matriz de confusão para o modelo Resnet, que pode ser usada para calcular várias métricas de desempenho, como precisão, recall e pontuação F1.

Continua na próxima página

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Deep Learning Applications for Dyslexia Prediction	Redes neurais convolucionais (CNNs), máquinas de Vetores de Suporte (SVMs), redes neurais artificiais (RNAs) e eliminação de características recursivas com validação cruzada (RFE-CV).	A aprendizagem por transferência foi aplicada no estudo sobre reconhecimento de caligrafia na dislexia. Especificamente, os autores desenvolveram um modelo de Rede Neural Convolucional (CNN) usando aprendizagem por transferência para esta tarefa.	O artigo menciona que os autores usaram um modelo CNN pré-treinado chamado VGG16 para aprendizagem por transferência em seu estudo sobre reconhecimento de escrita manual por dislexia.	De acordo com, o modelo CNN desenvolvido usando aprendizagem por transferência alcançou uma precisão de 94,5% na classificação de amostras de caligrafia disléxica, que superou outros modelos de aprendizagem de máquina usados no estudo. Os autores também relataram uma sensibilidade de 95,3% e uma especificidade de 93,7% para o modelo, indicando sua capacidade de identificar corretamente amostras de caligrafia disléxica e não disléxica.	Os autores usaram precisão, sensibilidade e especificidade para avaliar o desempenho do modelo CNN desenvolvido usando aprendizagem por transferência para reconhecimento de caligrafia em dislexia. O modelo alcançou precisão de 94,5%, sensibilidade de 95,3% e especificidade de 93,7%. Essas medidas são comumente usadas em aprendizado de máquina para avaliar o desempenho de um modelo na classificação correta de amostras de diferentes classes.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Recognition of Student Engagement State in a Classroom Environment Using Deep and Efficient Transfer Learning Algorithm	O estudo incorporou uma variante da Rede Neural Convolucional (CNN) para construir o modelo para a tarefa subjacente. Os pesquisadores treinaram os modelos Inceptionv3 e VGG16 em seu conjunto de dados coletados, mas o VGG16 oferece melhores resultados adicionando camadas densas estendidas ao conjunto de dados testado.	A aprendizagem por transferência foi aplicada no estudo usando um modelo VGG-16 previamente treinado, que foi ajustado para se adaptar à tarefa subsequente de reconhecer o estado de envolvimento do aluno.	O estudo utilizou a técnica de aprendizagem por transferência com um modelo VGG16 pré-treinado, que foi ajustado adicionando ou substituindo as camadas finais para corresponder ao número de turmas (engajadas e não engajadas) e adicionando camadas densas.	O modelo proposto prevê os frames engajados e não engajados com 90% de precisão e classificou 93% dos frames na classe positiva (recall); 93% dos frames foram classificados corretamente nas classes positiva e negativa (precisão).	As medidas de correção incluindo Precisão, Recall e Medida F foram utilizadas para o teste do modelo proposto. Essas medidas de acerto foram utilizadas para o cálculo dos estados afetivos dos alunos em um ambiente real de sala de aula.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
A Survey on Recent Approaches to Question Difficulty Estimation from Text	A pesquisa abrange uma ampla gama de abordagens, incluindo, entre outras, Máquinas de Vetores de Suporte (SVMs), Modelos de Mistura Gaussiana (GMMs), Redes Neurais Convolucionais (CNNs) e Aprendizagem por Transferência.	Os autores experimentaram três configurações diferentes de ELMo (pequeno, médio e original) e várias configurações de entrada (somente haste, apenas opções e haste e opções).	As técnicas utilizadas no processo de aprendizagem por transferência incluem o pré-treinamento de uma rede ELMo em um grande corpus de texto e o uso de um BiLSTM para aprender as informações sequenciais dos embeddings ELMo.	Os resultados indicam que a aprendizagem por transferência pode ser aplicada para melhorar a previsão da dificuldade da pergunta quando o tempo de resposta é usado como pré-treinamento, e a dificuldade é melhor prevista quando se usa apenas a haste do item. Contudo, o artigo não fornece métricas de desempenho específicas para o modelo.	Os autores experimentaram com três configurações ELMo diferentes (pequena, média e original) e várias configurações de entrada (apenas o caule, apenas as opções e o caule e as opções).

Continua na próxima página

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Improving the performance of automatic short answer grading using transfer learning and Augmentation	O estudo utiliza uma solução baseada em transfer learning com um modelo de transformer de sentenças pré-treinado e técnicas de aumento de dados, como substituição de sinônimos e troca aleatória de palavras. O estudo avalia o impacto do tamanho dos dados de treinamento no desempenho do modelo e compara várias combinações de técnicas de aumento de dados.	O transfer learning foi aplicado no estudo por meio do uso de modelos de transformer de sentenças pré-treinados, que foram finetuned no conjunto de dados SPRAG para a tarefa de avaliação automática de respostas curtas. O estudo também comparou o desempenho dos modelos finetuned com o desempenho de um modelo de referência baseado em BoW e TF-IDF.	O estudo utilizou três modelos de transformer de sentenças pré-treinados diferentes e avaliou o desempenho de cada modelo em diferentes tamanhos de dados de treinamento e combinações de técnicas de aumento de dados. O estudo também comparou o desempenho dos modelos finetuned com o desempenho de um modelo de referência baseado em BoW e TF-IDF. As técnicas de aumento de dados utilizadas no estudo incluem substituição de sinônimos e troca aleatória de palavras.	O desempenho do modelo foi avaliado em diferentes tamanhos de dados de treinamento e combinações de técnicas de aumento de dados. O estudo apresenta tabelas com os resultados de precisão e F1-score para cada modelo treinado e para cada tamanho de dados de treinamento. Os resultados mostram que o desempenho dos modelos finetuned é melhor do que o modelo de referência baseado em BoW e TF-IDF. Além disso, o estudo mostra que o desempenho do modelo é melhorado com o aumento do tamanho dos dados de treinamento e com o uso de técnicas de aumento de dados, como substituição de sinônimos e troca aleatória de palavras.	Medidas de precisão e F1-score para a tarefa de classificação e coeficientes de correlação de Pearson e Spearman para a tarefa de similaridade.

Continua na próxima página

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Automatic engagement estimation in smart education/learning settings: a systematic review of engagement definitions, datasets, and methods	O artigo discute vários métodos baseados em aprendizado de máquina usados na literatura para estimativa automática de engajamento, incluindo algoritmos clássicos de aprendizado de máquina, como máquinas de vetores de suporte (SVMs) e árvores de decisão (DTs), bem como algoritmos de aprendizado profundo.	A aprendizagem por transferência envolve o uso de uma rede neural pré-treinada em um grande conjunto de dados para extrair recursos para usar em tarefas com conjuntos de dados menores. O artigo fornece uma lista de conjuntos de dados de imagens usados para aprendizagem por transferência, incluindo FER-2013, VGG-Face, VGGFace2, FaceNet, AffectNet, 300W-LP e AFLW2000, JAFFE, CK+ e RAF-DB.	O artigo não fornece informações específicas sobre as técnicas utilizadas no processo de aprendizagem por transferência.	Os resultados de desempenho variam dependendo do conjunto de dados, método e métrica de avaliação utilizados em cada estudo, e o artigo não fornece uma lista abrangente de todos os resultados de desempenho na literatura.	Exatidão, precisão, recall, pontuação F1, área sob a curva (AUC), média absoluta erro (MAE), raiz do erro quadrático médio (RMSE) e muito mais.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Hand Gesture Recognition with ConvNets for School-Aged Children to Learn Basic Arithmetic Operations	Deep Learning (DL), Transfer Learning (TL) e Redes Neurais (CNN).	A técnica Transfer Learning (TL), foi selecionada para esta tarefa, pois foi usada anteriormente para gestos manuais e reconhecimento. Em termos gerais, a ideia principal do TL é ter um modelo que foi pré-treinado em um ótimo número de imagens, um modelo significativamente bem treinado é alcançado sem uma grande amostra para fins de reciclagem.	Foi utilizado o modelo que foi pré-treinado/Pré-processamento em um ótimo número de imagens.	O processo de validação resultou em uma precisão de 88,1%, bem como uma perda total de 0,49.	Acurácia, Precisão Média, Recuperação Média, Perda de Classificação, Perda de localização, Perda de regularização e Perda total.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Bimodal Learning Engagement Recognition from Videos in the Classroom	O estudo mencionado que comparou sete classificadores, que vão desde modelos mais simples, como árvores de decisão, até modelos mais complexos, incluindo máquinas de vetores de suporte (SVM). Os resultados indicam que o algoritmo KNN pode classificar bem o envolvimento na aprendizagem bimodal.	O método de aprendizagem por transferência foi aplicado no canal de envolvimento emocional usando ResNet50 como modelo pré-treinado e ajustando-o para reconhecer o envolvimento emocional dos alunos. No canal de envolvimento comportamental, o conjunto de dados de envolvimento comportamental autoconstruído foi usado para treinar a rede CoAtNet para estimar o envolvimento comportamental dos alunos.	O método de aprendizagem por transferência foi usado no canal de engajamento emocional, aplicando o ResNet50 como modelo pré-treinado e ajustando-o para reconhecer o envolvimento emocional do aluno.	O modelo alcançou alta precisão, com cada índice atingindo mais de 95%, o que atende aos requisitos de precisão de detecção de alunos em sala de aula. Especificamente, em comparação com a perda de IoU existente e as perdas baseadas em IoU, incluindo GIoU, DIoU e CIoU, YOLOv5 com perda de α -IoU obteve os resultados de detecção mais eficazes, melhorando a precisão, recall, AP, mAP@0,5 e mAP@. 5: 0,95 valores por 1,4%, 1,8%, 0,6%, 0,9% e 0,4%, respectivamente.	Precisão, recall, AP, mAP@.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Development of Low Resource Machine Learning Models for Child Cognitive Ability Assessments	O estudo utilizou vários tipos de soluções para diferentes tarefas. Para classificação de imagens, o estudo utilizou modelos CNN, incluindo aprendizagem por transferência com formas desenhadas à mão e formas puramente geradas por computador. Para reconhecimento de expressões faciais, o estudo utilizou um modelo híbrido combinando CNN e SVM.	O Transfer Learning foi aplicado para a classificação de formas desenhadas à mão. O modelo foi treinado inicialmente com um conjunto de dados de formas geométricas 2D e, em seguida, transferido para um conjunto de dados de 2000 imagens de formas desenhadas à mão por crianças, que pertenciam a 3 classes: triângulo, círculo e quadrado. Além disso, a técnica de aumento de dados (data augmentation) foi aplicada a esse conjunto de dados para melhorar o desempenho do modelo.	O estudo utilizou a técnica de Transfer Learning para a classificação de formas desenhadas à mão. Além disso, a técnica de aumento de dados (data augmentation) foi aplicada a esse conjunto de dados para melhorar o desempenho do modelo.	O modelo CNN usado para classificação de letras alcançou 97% de precisão para o conjunto de dados contendo letras cingalesas desenhadas à mão e geradas. O modelo de classificação de formas alcançou 99% de precisão para o conjunto de dados contendo formas puramente geradas por computador. O modelo híbrido de aprendizagem profunda superou outros modelos com uma pontuação de precisão de 89,1% e uma pontuação de recall de 88,6%. O modelo LVCSR para respostas verbais de crianças cingalesas alcançou quase 40,1% usando a abordagem estatística ASR HMM-GMM. O modelo alcançou uma pontuação de precisão de 89,1% e uma pontuação de recall de 88,6%.	O estudo usou pontuações de precisão e recall para avaliar o desempenho do modelo de aprendizagem por transferência para identificar formas desenhadas à mão.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Deep Learning Based Emotion Recognition and Visualization of Figural Representation	O estudo propõe uma abordagem baseada em aprendizagem profunda usando o algoritmo CNN-BiLSTM para reconhecimento de emoções a partir de imagens de expressões faciais e dados de voz coletados durante cursos online.	O transfer learning foi aplicado no estudo por meio da utilização de um modelo de aprendizado profundo pré-treinado para reconhecimento de emoções em imagens.	O estudo utilizou a técnica de transferência de aprendizado fina, que envolveu o ajuste do modelo pré-treinado VGG-Face para o reconhecimento de emoções em imagens de expressões faciais de alunos. Além disso, o estudo também utilizou a técnica de aumento de dados para aumentar o tamanho do conjunto de dados de treinamento e melhorar o desempenho do modelo.	O modelo proposto alcançou uma acurácia de reconhecimento de emoções de 87,5%, que é significativamente maior do que a acurácia alcançada pelos outros modelos.	O desempenho do modelo de aprendizagem por transferência foi avaliado usando uma série de métricas de avaliação, incluindo exatidão, precisão, recall e pontuação F1.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Meta Transfer Learning for Early Success Prediction In MOOCs	O artigo descreve o uso da aprendizagem por meta-transferência, que combina dados de interação e informações do curso para prever o sucesso dos alunos em MOOCs.	Para aplicar a aprendizagem por transferência neste contexto, o artigo propõe uma abordagem de meta-aprendizagem que combina dados de interação e informações do curso de vários MOOCs para criar um modelo generalizável para previsão de sucesso precoce.	O artigo apresenta três estratégias para criar modelos generalizáveis para a tarefa de predição de sucesso precoce em MOOCs usando transfer learning. No entanto, o texto não especifica quais técnicas de transfer learning foram utilizadas.	O desempenho do modelo varia dependendo do contexto e das estratégias utilizadas. Em média, a taxa de transferência de comportamento é de 75% para cursos com um número razoavelmente alto de quizzes e 57% para cursos com um número baixo de quizzes, em ambos os níveis de predição precoce. Para cursos nunca vistos durante o treinamento, a taxa de transferência de comportamento é de cerca de 70% em ambos os níveis de predição precoce.	Acurácia, a área sob a curva ROC (AUC-ROC), a precisão, a revocação e a medida F1. Além disso, o artigo também apresenta a taxa de transferência de comportamento (BAC), que é uma medida de quão bem o modelo é capaz de transferir o comportamento de um conjunto de estudantes para outro conjunto de estudantes.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Classification in Math Class: Using Convolutional Neural Networks to Categorize Student Cognitive Demand	O estudo utilizou uma solução que envolveu transfer learning com uma rede neural convolucional (CNN) chamada MobileNet V2 e um modelo de máquina de vetores de suporte (SVM) para classificação.	O transfer learning foi aplicado no estudo utilizando a rede neural convolucional MobileNet V2, que foi pré-treinada em um grande conjunto de dados de imagens chamado ImageNet. Em seguida, a rede foi ajustada (fine-tuned) para a tarefa específica de classificação de demanda cognitiva dos alunos em sala de aula.	O artigo não menciona especificamente quais técnicas foram utilizadas no processo de transfer learning. No entanto, é mencionado que a rede neural convolucional MobileNet V2 foi utilizada como base para o modelo e que a transferência de aprendizado foi aplicada aos parâmetros treináveis da camada Dense Bottleneck da rede.	É mencionado que os modelos foram treinados até atingirem uma precisão de treinamento superior a 85%, um valor semelhante à precisão humana aplicada às etiquetas originais de demanda cognitiva.	O artigo não fornece uma resposta clara sobre as medidas específicas usadas para avaliar o desempenho do modelo de transfer learning. No entanto, é mencionado que o modelo experimental foi comparado com o modelo de demanda cognitiva de linha de base para avaliar o desempenho entre os dois modelos.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Generating Answer Candidates for Quizzes and Answer-Aware Question Generators	O estudo propõe uma abordagem que combina informações ortográficas, lexicais, sintáticas e semânticas para gerar respostas candidatas que possam servir como entrada para modelos de geração de perguntas conscientes de respostas. O estudo menciona que a abordagem proposta é competitiva em relação a modelos neurais complexos e apresenta melhorias em relação a abordagens simples baseadas em entidades nomeadas.	O artigo discute o pré-treinamento de um modelo de linguagem unificado para compreensão e geração de linguagem natural. O transfer learning foi aplicado neste estudo por meio do pré-treinamento de um modelo de linguagem unificado para a tarefa de geração de perguntas e respostas.	O estudo não se concentra em técnicas de transfer learning específicas. Embora mencione o sucesso de grandes modelos pré-treinados, como BERT, RoBERTa, ALBERT, T5 e BART, em tarefas de geração de perguntas. Em vez disso, a abordagem proposta combina várias informações linguísticas para gerar respostas candidatas que possam servir como entrada para modelos de geração de perguntas conscientes de respostas.	O estudo menciona que a abordagem proposta é competitiva em relação a modelos neurais complexos e apresenta melhorias em relação a abordagens simples baseadas em entidades nomeadas. O estudo não fornece informações específicas sobre o desempenho do modelo em termos de métricas de avaliação, como precisão, recall ou F1-score.	O estudo não fornece informações específicas sobre as medidas usadas para avaliar o desempenho do modelo.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
A deep transfer learning approach to modeling teacher discourse in the Classroom	A primeira abordagem é uma abordagem tradicional de aprendizado de máquina que usa uma combinação de redes neurais convolucionais (CNNs) e máquinas de vetores de suporte (SVMs) para classificar características do discurso dos professores. A segunda abordagem é uma abordagem de aprendizagem por transferência profunda que aproveita o poder dos modelos baseados em transformadores, especificamente o modelo BERT para classificar características do discurso do professor.	Para aplicar a aprendizagem por transferência, os autores usaram um modelo BERT pré-treinado que foi treinado em um grande corpus de dados de texto. Eles então ajustaram o modelo BERT pré-treinado na tarefa específica de classificar características do discurso do professor usando um conjunto de dados menor de falas anotadas do professor.	Os autores usaram um modelo BERT (Bidirecional Encoder Representations from Transformers) pré-treinado para aprendizagem por transferência. Os autores também usaram uma técnica chamada parada antecipada para evitar overfitting durante o ajuste fino. A parada antecipada envolve monitorar o desempenho do modelo em um conjunto de validação durante o treinamento e interromper o processo de treinamento quando o desempenho no conjunto de validação começar a diminuir.	O modelo BERT alcançou uma precisão média de 0,82 e uma pontuação F1 média de 0,77 em todos os sete recursos, enquanto a abordagem tradicional de aprendizado de máquina alcançou uma precisão média de 0,75 e uma pontuação F1 média de 0,68.	Os autores usaram várias medidas, incluindo exatidão, pontuação F1, precisão e recall.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Modeling review helpfulness with augmented transformer neural Networks	O artigo discute vários tipos de modelos de aprendizado de máquina usados na pesquisa, incluindo modelos clássicos, como regressão logística, floresta aleatória, vários sabores de modelos Bayesianos ingênuos e máquinas de vetores de suporte, bem como modelos de redes neurais, como text-CNN, LSTM e um modelo baseado em transformador.	O transfer learning foi aplicado no estudo para reduzir o número de parâmetros treináveis em modelos de transformadores, a fim de melhorar o desempenho do modelo.	No processo de transfer learning, o estudo utilizou o modelo BERT (Bidirectional Encoder Representations from Transformers) pré-treinado em grandes conjuntos de dados. Em seguida, o modelo foi ajustado para a tarefa específica de detecção de utilidade de feedback. Durante o treinamento do modelo, o tokenizador foi modificado simultaneamente para adotar um novo vocabulário, fornecendo um novo tokenizador quando o modelo é treinado.	O desempenho do modelo foi avaliado em comparação com outros modelos, utilizando a métrica F1 score. O modelo resultante, que utilizou transfer learning com o modelo BERT, obteve um F1 score de 0,68 na tarefa de detecção de utilidade de feedback.	O desempenho do modelo de transfer learning foi avaliado utilizando a métrica F1 score, que é uma medida de precisão e recall combinados.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Personalized Online Education Learning Strategies Based on Transfer Learning Emotion Classification Model	Aprendizagem por transferência e CNN.	No estudo, o transfer learning foi aplicado para treinar um conjunto de dados de ação experimental usando um grande conjunto de dados regional do conjunto de dados de vídeo MOOC. O modelo proposto captura com sucesso o vetor de características para a primeira tarefa e, em seguida, redefine a função de convolução com uma camada totalmente conectada adicional e retrai o vetor de características.	No processo de transfer learning, foram utilizadas técnicas de extração de características baseadas em seleção livre e redes neurais convolucionais (CNNs) para reconhecimento de comportamento humano. O modelo proposto utiliza a arquitetura C3D para extrair características de quatro comportamentos diferentes dos alunos na sala de aula: entrar, sentar, ficar em pé e sair da sala.	Os resultados experimentais mostram que, no processo de “mineração massiva de dados”, a precisão deste modelo é 90,30% melhor que a linha de base.	O artigo não fornece informações sobre outras medidas utilizadas para avaliar o desempenho do modelo de aprendizagem por transferência.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Math Word Problem Solver Based On Text-to-Text Transformer Model	O estudo propõe um modelo de aprendizagem profunda baseado no modelo Text-to-Text Transformer (T5), que é um modelo de linguagem pré-treinado em larga escala. O modelo T5 é ajustado em uma grande quantidade de dados e então usado para resolver tarefas posteriores, como problemas matemáticos.	O transfer learning foi aplicado por meio da técnica de pré-treinamento em uma grande quantidade de dados e, em seguida, ajustando o modelo para a tarefa específica de resolução de problemas matemáticos em linguagem natural. O modelo foi pré-treinado em um corpus chinês usando uma técnica de modelagem de linguagem natural e, em seguida, ajustado para a tarefa de resolução de problemas matemáticos.	No processo de transfer learning, foram utilizadas técnicas de pré-treinamento em um corpus chinês usando uma técnica de modelagem de linguagem natural, que é uma técnica comum em transfer learning. Além disso, o modelo foi ajustado para a tarefa de resolução de problemas matemáticos em linguagem natural usando três estratégias: substituição aleatória de palavras de entrada, penalidade de gradiente e mecanismo de replicação.	O modelo proposto no estudo alcançou uma precisão geral de 78,61% na resolução de problemas matemáticos em chinês. Além disso, os problemas foram classificados em seis categorias com base nos pontos de conhecimento envolvidos, e o estudo explora o desempenho do modelo em diferentes tipos de problemas.	O estudo utilizou a precisão como principal medida para avaliar o desempenho do modelo de aprendizagem por transferência.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Development of CNN Transfer Learning for Dyslexia Handwriting Recognition	O estudo se concentrou no desenvolvimento de uma solução de transfer learning para reconhecimento de caligrafia disléxica usando uma Rede Neural Convolucional (CNN) baseada na arquitetura famosa de reconhecimento de caligrafia usando LeNet-5.	O transfer learning foi aplicado usando a técnica de fine-tuning, que envolveu o ajuste dos pesos de uma rede neural pré-treinada (LeNet-5) em um novo conjunto de dados (dataset de caligrafia disléxica). Além disso, foram utilizadas técnicas de aumento de dados para aumentar a quantidade de dados de treinamento disponíveis e melhorar o desempenho do modelo.	Fine-tuning para ajuste dos pesos de uma rede neural pré-treinada (LeNet-5) em um novo conjunto de dados (dataset de caligrafia disléxica), inicialização de pesos da rede pré-treinada com pesos treinados em um grande conjunto de dados de imagens (MNIST) e aumento de dados para aumentar a quantidade de dados de treinamento disponíveis e melhorar o desempenho do modelo.	A acurácia de classificação para as classes "Normal", "Reversal" e "Corrected" foram de 0,92, 0,97 e 0,97, respectivamente. O modelo alcançou uma acurácia notável de 95,34% na classificação dessas três classes. Além disso, o estudo também menciona que o modelo foi capaz de alcançar alta acurácia de treinamento e validação, com uma acurácia de treinamento de 0,9823 e uma acurácia de validação de 0,9702, usando um valor de dropout de 0,1.	Métricas de desempenho, como acurácia de treinamento e validação, perda de treinamento e validação, tempo de treinamento e teste, também foram avaliadas.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
DRAFT: A Novel Framework to Reduce Domain Shifting in Self-supervised Learning and Its Application to Children's ASR	A estrutura é avaliada com três abordagens amplamente utilizadas, incluindo aprendizagem por transferência com CNN e SVM, e mostra resultados promissores em tarefas de ASR para crianças. No entanto, os detalhes específicos da abordagem de aprendizagem por transferência com CNN e SVM não são discutidos neste artigo.	O artigo não fornece detalhes específicos sobre como a aprendizagem por transferência com CNN e SVM foi aplicada no estudo. A estrutura proposta, DRAFT, foi projetada para reduzir a mudança de domínio em modelos de fala pré-treinados por meio de um estágio de adaptação adicional e é avaliada com três abordagens amplamente utilizadas, incluindo APC, Wav2vec2.0 e HuBERT.	Os detalhes específicos da aprendizagem por transferência com CNN e SVM não são discutidos neste artigo.	A estrutura DRAFT proposta alcançou melhorias relativas de WER de até 19,7% nas duas tarefas de ASR infantil quando comparada com as linhas de base de ajuste fino convencionais sem adaptação.	As medidas específicas utilizadas para transferência de aprendizagem com CNN e SVM não são discutidas neste artigo.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
EXAMS: A Multi-Subject High School Examinations Dataset for Cross-Lingual and Multilingual Question Answering	O estudo utilizou transfer learning com modelos pré-treinados multilíngues, como mBERT e XLM-RoBERTa.	O transfer learning foi aplicado no estudo por meio do uso de modelos pré-treinados multilíngues, como mBERT e XLM-RoBERTa, que foram ajustados para prever a resposta correta em um ambiente de múltipla escolha, dada uma seleção de contexto.	O estudo utilizou o método de fine-tuning, que é uma técnica comum de transfer learning em que um modelo pré-treinado é ajustado para uma tarefa específica. Além disso, o estudo utilizou modelos pré-treinados multilíngues, como mBERT e XLM-RoBERTa, que foram treinados em grandes conjuntos de dados multilíngues para aprender representações de linguagem geral que podem ser transferidas para tarefas específicas em diferentes idiomas.	Os resultados mostraram que os modelos pré-treinados multilíngues tiveram um desempenho significativamente melhor do que os modelos treinados a partir do zero em todos os conjuntos de dados avaliados, incluindo o Exams. Por exemplo, o modelo XLM-RoBERTa obteve uma pontuação de 70,5% no conjunto de dados Exams, enquanto o modelo treinado a partir do zero obteve uma pontuação de apenas 38,5%.	Para avaliar o desempenho dos modelos de transfer learning, o estudo utilizou várias medidas de avaliação comuns em tarefas de perguntas e respostas, como a precisão (accuracy), que mede a proporção de perguntas respondidas corretamente pelo modelo, e a F1-score, que é uma medida que combina a precisão e a cobertura (recall) do modelo.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Spontaneous facial expression database for academic emotion Inference in online learning	Em relação à solução utilizada no estudo, os autores empregaram aprendizagem por transferência com um modelo de rede neural convolucional (CNN) VGG16 pré-treinado. O modelo pré-treinado do ImageNet foi selecionado para inicializar o CNN, e apenas as camadas mais profundas que CONV3-1 (incluem) são adaptadas no OL-SFED.	Os autores empregaram um modelo de rede neural convolucional (CNN) VGG16 pré-treinado, que foi inicializado com pesos do conjunto de dados ImageNet. Apenas as camadas mais profundas que CONV3-1 (incluem) foram adaptadas no OL-SFED usando aprendizagem por transferência.	No processo de transfer learning, foi utilizada a técnica de inicialização do modelo CNN com o modelo pré-treinado do ImageNet e a adaptação de camadas específicas do modelo VGG16 treinado no ImageNet para o OL-SFED.	O modelo proposto alcançou uma acurácia média de 85,5% na classificação das emoções acadêmicas em imagens e 81,3% na classificação em vídeos.	Foram utilizadas várias medidas para avaliar o desempenho do modelo de transfer learning, incluindo recall, precisão, F1, acurácia média e o coeficiente kappa (k).
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Relationships among cognitive components in Logo learning and Transfer.	O estudo utilizou uma abordagem componencial para modelar relações preditivas entre os componentes cognitivos da memória de trabalho, representações, metacognição e desempenho, e uma série de tarefas de transferência.	O estudo aplicou o conceito de transferência de aprendizagem (transfer learning) para avaliar a eficácia do ensino de programação em Logo. Foram desenvolvidas várias medidas de transferência que envolviam os componentes cognitivos avaliados no estudo e que pesquisas anteriores haviam indicado como elementos transferíveis da experiência com Logo. O estudo também usou modelagem de equações estruturais e análises de caminho para identificar e avaliar as relações preditivas diretas e indiretas entre os componentes cognitivos hipotetizados e as tarefas de transferência.	O estudo não especifica quais técnicas de transfer learning foram utilizadas no processo de ensino de programação em Logo. Em vez disso, o estudo se concentra em avaliar a eficácia do ensino de programação em Logo e como diferentes fatores cognitivos e contextuais afetam a transferência de habilidades.	O estudo constatou que os resultados fornecem um forte suporte para as expectativas originais relativas aos padrões de relações entre componentes cognitivos, desempenho do Logo e desempenho em tarefas de transferência. Os dados sugerem algumas pequenas modificações do modelo original, mas no geral, o modelo foi eficaz na identificação de relações entre as variáveis.	O estudo utilizou várias medidas para avaliar o desempenho e a transferência do Logo, incluindo avaliações de previsão de comando de posição/direção, conhecimento de como usar variáveis, domínio de comando, domínio de previsão, domínio de produção, analogias, detecção de erros, monitoramento de compreensão, conhecimento matemático e planejamento. O estudo utilizou então a análise de caminhos para explorar padrões de relações entre os componentes cognitivos e medidas de transferência.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Learning and Transfer Effects of Embodied Simulations Targeting Cross-cutting Concepts in Science	Não há menção a uma solução específica de aprendizado de máquina neste artigo. O estudo se concentra em como as simulações corporificadas podem melhorar a compreensão dos alunos sobre conceitos científicos complexos.	Explora a extensão em que os indivíduos que adquirem conhecimento acadêmico por meio de uma experiência corporificada são capazes de transferir esse conhecimento para contextos novos.	O artigo não descreve técnicas específicas utilizadas no processo de aprendizagem por transferência. Em vez disso, ele apresenta dois experimentos que exploram como atividades físicas podem melhorar a compreensão dos alunos sobre conceitos científicos complexos.	O artigo não apresenta um modelo específico para desempenho. Em vez disso, ele apresenta os resultados de dois experimentos que exploram como atividades físicas podem melhorar a compreensão dos alunos sobre conceitos científicos complexos.	As medidas usadas para avaliar o desempenho dos alunos incluem testes de pré e pós-teste, bem como testes de transferência de conhecimento para novos contextos. O artigo discute as análises estatísticas realizadas para avaliar a eficácia das simulações corporificadas em melhorar o desempenho dos alunos nessas medidas.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Interdomain Transfer Between Isomorphic Topics in Algebra And Physics	O foco do estudo está nos experimentos conduzidos por pesquisadores para investigar a transferência de resolução de problemas entre dois subdomínios da álgebra e da física.	O foco do artigo está nos experimentos conduzidos por pesquisadores para investigar a transferência de resolução de problemas entre dois subdomínios da álgebra e da física. Os pesquisadores mediram a transferência avaliando os métodos de solução aplicados aos problemas, em vez das respostas corretas finais. Eles categorizaram os métodos de solução exibidos nos trabalhos escritos dos sujeitos em duas categorias principais, o método “aprendido” e os “outros” métodos.	O artigo não fornece informações sobre a aplicação da aprendizagem por transferência no estudo.	O artigo não fornece informações sobre o desempenho do modelo.	Os pesquisadores mediram a transferência avaliando os métodos de solução aplicados aos problemas, em vez das respostas corretas finais. A principal medida dependente de transferência no presente estudo foi se o método aprendido foi aplicado a problemas estruturalmente isomórficos, mas desconhecidos.
Continua na próxima página					

TÍTULO	TIPO DE SOLUÇÃO	COMO O TRANSFER LEARNING FOI APLICADO NO ESTUDO?	QUAIS TÉCNICAS FORAM UTILIZADAS NO PROCESSO DE TRANSFER LEARNING?	QUAL DESEMPENHO DO MODELO?	QUAIS MEDIDAS FORAM USADAS PARA AVALIAR O DESEMPENHO DO MODELO DE TRANSFER LEARNING?
Depth and Breadth of Vocabulary in Two Languages: Which Vocabulary Skills Transfer?	O estudo examinou a transferência de habilidades de vocabulário entre duas línguas em alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos, por meio de testes de vocabulário em inglês e espanhol, bem como de uma tarefa de definição e descrição de palavras em ambas as línguas.	O estudo examinou a transferência de habilidades de vocabulário entre duas línguas em alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos, por meio de testes de vocabulário em inglês e espanhol, bem como de uma tarefa de definição e descrição de palavras em ambas as línguas.	O estudo não utilizou técnicas específicas de transfer learning. Em vez disso, o estudo examinou a transferência de habilidades de vocabulário entre duas línguas em alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos, por meio de testes de vocabulário em inglês e espanhol, bem como de uma tarefa de definição e descrição de palavras em ambas as línguas.	Neste estudo não foram utilizadas métricas específicas para avaliar o desempenho do modelo.	Não foram utilizadas medidas específicas para avaliar o desempenho do modelo. Em vez disso, o estudo examinou a transferência de habilidades de vocabulário entre duas línguas em alunos bilíngues da 4ª e 5ª séries em escolas primárias públicas nos Estados Unidos, por meio de testes de vocabulário em inglês e espanhol, bem como de uma tarefa de definição e descrição de palavras em ambas as línguas.

Fonte: Autoria Própria (2024).