



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE ENGENHARIAS E TECNOLOGIA
CURSO INTERDISCIPLINAR EM TECNOLOGIA DA INFORMAÇÃO

DANIEL LINS SILVA ANDRADE

**PREDIÇÃO DE ÍNDICES PSICOLÓGICOS EM AMBIENTES DIGITAIS
UTILIZANDO REDES NEURAIS PROFUNDAS**

PAU DOS FERROS

2025

DANIEL LINS SILVA ANDRADE

**PREDIÇÃO DE ÍNDICES PSICOLÓGICOS EM AMBIENTES DIGITAIS
UTILIZANDO REDES NEURAIS PROFUNDAS**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito parcial para obtenção do título de Bacharel Interdisciplinar em Tecnologia da Informação.

Orientador: George Felipe Fernandes Vieira, Prof. Bel.

Coorientadora: Náthalee Cavalcanti de Almeida Lima, Profa. Dra.

PAU DOS FERROS

2025

© 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.

A553p Andrade, Daniel Lins Silva.
Predição de índices psicológicos em ambientes
digitais utilizando redes neurais profundas /
Daniel Lins Silva Andrade. - 2025.
55 f. : il.

Orientador: George Felipe Fernandes Vieira.
Coorientador: Náthalee Cavalcanti de Almeida
Lima.

Monografia (graduação) - Universidade Federal
Rural do Semi-árido, Curso de Tecnologia da
Informação, 2025.

1. Redes neurais. 2. Predição de ansiedade. 3.
Fobia social. 4. Satisfação com a vida. 5. Dados
de jogos. I. Vieira, George Felipe Fernandes,
orient. II. Lima, Náthalee Cavalcanti de Almeida,
co-orient. III. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade
com AACR2 e os dados fornecidos pelo(a) autor(a).

Biblioteca Campus Pau dos Ferros
Bibliotecária: Erlanda Maria Lopes da Silva
CRB: 15/966

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.

DANIEL LINS SILVA ANDRADE

**PREDIÇÃO DE ÍNDICES PSICOLÓGICOS EM AMBIENTES DIGITAIS
UTILIZANDO REDES NEURAIS PROFUNDAS**

Monografia apresentada à Universidade Federal Rural do Semi-Árido como requisito parcial para obtenção do título de Bacharel Interdisciplinar em Tecnologia da Informação.

Defendida em: 05 / 12 / 2025

BANCA EXAMINADORA

George Felipe Fernandes Vieira, Prof. Bel. (UFERSA)
Presidente

Náthalee Cavalcanti de Almeida Lima, Profa. Dra. (UFERSA)
Membro Examinador

João Batista de Souza Neto, Prof. Dr. (UFERSA)
Membro Examinador

Maria Camilla Souza Trindade, Psic. Me. (UFERSA)
Membro Examinador

Dedico este trabalho à Deus, pois com sua graça eu cheguei até aqui, e pela honra e glória d'Ele, eu concluo esta trajetória. Aos meus pais, sangue do meu sangue, que foram minhas pernas quando não pude andar e meus braços quando não pude avançar. Aos meus amigos, que nos momentos alegres e tristes, confiaram e lutaram ao meu lado.

AGRADECIMENTOS

Agradeço antes de tudo, à Deus, que me proporcionou não apenas a oportunidade da graduação, mas mudar de vida, crescer e amadurecer. Foram muitas madrugadas, muitos dias difíceis, e os desafios só aumentaram, mas cada oração, cada gota de lágrima e sangue derramada, foram honradas e após tanta luta, pude ser agraciado com o fim dessa jornada.

Agradeço também aos meus pais, Linsmar e Elisangela, que intercederam por mim, do começo ao fim, contra todos os males da vida, estiveram de pé e venceram juntos comigo.

Ao meu orientador Prof. George Felipe, que assim como um dos vários milagres de Deus, entrou em minha vida e me proporcionou não só o conhecimento, mas batalhou junto comigo para seguir com a proposta deste trabalho, nas madrugadas, nas tardes, no almoço, janta e em vários momentos, seguiu lutando comigo.

À minha coorientadora Profa. Náthalee Cavalcanti, que me direcionou da melhor maneira, acreditou no meu projeto e aceitou a proposta, soube ouvir, soube guiar e teve além de empatia, humanidade ao me conduzir neste trabalho.

Agradeço aos vários amigos que atravessaram juntos comigo essa história, que me ensinaram de diversas formas e puderam me tornar alguém melhor, tanto como acadêmico, quanto como pessoa.

À minha amiga Alana, que confiou desde sempre que este era o caminho certo e esteve ao meu lado. Ao meu amigo Gerson, que esteve ao meu lado nos momentos que até mesmo meu corpo e minha saúde me abandonaram. A minha amiga Daniele, que como uma brisa leve neste ambiente árido de clima e convivência, me trouxe calma e paz em vários momentos, me ensinando a lidar com as adversidades da vida, sempre sendo resiliente. Aos meus amigos Leonardo e Ricardo, que se tornaram meus irmãos com o passar dos anos, me ensinaram, brigaram e lutaram comigo, mostrando como a vida pode ser difícil mas que eu definitivamente não estava só. A minha amiga Maria, que sempre esteve ali para desabafar. A minha amiga Milena, que se encontrou comigo no começo e foi um apoio de força em muitos momentos. Ao meu amigo Kairo, que assim como a origem de seu nome, me mostrou que tudo é no tempo de Deus.

À todos amigos que aqui não foram citados, mas que passaram por minha vida e deixaram uma parte de si nesta história, meu maior respeito e agradecimento, vocês são parte de mim.

EPIÍGRAFE

“Até o dia em que Deus dignar-se a desvelar o futuro para o homem, toda a sabedoria humana estará nestas palavras: Esperar e ter esperança.”

(Alexandre Dumas, *O Conde de Monte Cristo*, ed. Zahar, 2008)

RESUMO

Os jogos digitais destacam-se como ambientes relevantes para pesquisas comportamentais devido à capacidade de estimular mecanismos de recompensa e interação social. Entretanto, a exposição contínua a esses sistemas pode reforçar comportamentos associados à ansiedade, impulsividade e fobia social. Este estudo aplica redes neurais densas (DNN) para prever níveis clínicos de ansiedade (GAD-7), fobia social (SPIN) e satisfação com a vida (SWL) a partir de variáveis comportamentais autodeclaradas. A escolha pelas DNN decorre de sua capacidade de modelar relações não lineares complexas entre atributos comportamentais e indicadores psicológicos, superando limitações de métodos tradicionais. Adotou-se uma abordagem quantitativa envolvendo pré-processamento, normalização, balanceamento de classes e modelagem com TensorFlow/Keras. Os modelos alcançaram acurácias de 81,9% (GAD-7), 81,7% (SPIN) e 74,9% (SWL), demonstrando estabilidade e interpretabilidade adequadas. A principal contribuição desta pesquisa consiste em integrar dados de comportamento digital com índices clínicos validados, preenchendo uma lacuna na literatura ao aplicar aprendizado profundo em cenários não clínicos, como o contexto de jogos online. Os resultados indicam o potencial das redes neurais como ferramenta complementar para avaliação psicológica preliminar e monitoramento de bem-estar em ambientes digitais.

Palavras-chave: redes neurais; predição de ansiedade; fobia social; satisfação com a vida; dados de jogos

ABSTRACT

Digital games stand out as relevant environments for behavioral research due to their ability to stimulate reward mechanisms and social interaction. However, continuous exposure to these systems can reinforce behaviors associated with anxiety, impulsivity, and social phobia. This study applies Deep Neural Networks (DNN) to predict clinical levels of anxiety (GAD-7), social phobia (SPIN), and life satisfaction (SWL) based on self-reported behavioral variables. The choice of DNN models is justified by their capacity to capture complex nonlinear relationships between behavioral attributes and psychological indicators, surpassing the limitations of traditional methods. A quantitative approach was adopted, including data preprocessing, normalization, class balancing, and predictive modeling using TensorFlow/Keras. The models achieved accuracies of 81.9% (GAD-7), 81.7% (SPIN), and 74.9% (SWL), demonstrating adequate stability and interpretability. The main contribution of this research lies in integrating digital behavior data with validated clinical indices, addressing a gap in the literature by applying deep learning in non-clinical contexts such as gaming environments. The findings highlight the potential of neural networks as a complementary tool for preliminary psychological assessment and for monitoring well-being in digital settings.

Keywords: neural networks; anxiety prediction; social phobia; life satisfaction; gaming data

LISTA DE FIGURAS

Figura 1 – <i>Heatmap</i> da coluna <i>Birthplace</i> após organização	32
Figura 2 – <i>Heatmap</i> da coluna <i>Residence</i> após organização	32
Figura 3 – Distribuição inicial dos níveis de ansiedade	35
Figura 4 – Distribuição após o balanceamento dos valores dos níveis de ansiedade . . .	36
Figura 5 – Distribuição da variável <i>stream</i> antes do tratamento	36
Figura 6 – Distribuição da variável <i>stream</i> depois do tratamento	36
Figura 7 – Distribuição da variável <i>Age</i> antes do tratamento	37
Figura 8 – Distribuição da variável <i>Age</i> depois do tratamento	37
Figura 9 – Distribuição da variável <i>Hours</i> antes do tratamento	37
Figura 10 – Distribuição da variável <i>Hours</i> depois do tratamento	37
Figura 11 – Histograma do índice clínico GAD-7	37
Figura 12 – Histograma do índice clínico SPIN	38
Figura 13 – Histograma do índice clínico SWL	38
Figura 14 – Arquitetura da Rede Neural Densa responsável pelo índice GAD-7	40
Figura 15 – Gráfico de acurácia da classificação por época para o índice GAD-7	42
Figura 16 – Gráfico de acurácia da classificação por época para o índice SWL	42
Figura 17 – Gráfico de acurácia da classificação por época para o índice SPIN	42
Figura 18 – Matriz de confusão para o índice GAD-7	44
Figura 19 – Matriz de confusão para o índice SPIN	44
Figura 20 – Matriz de confusão para o índice SWL	44
Figura 21 – Boxplot de relação dos gêneros com a média do índice GAD-7	45
Figura 22 – Boxplot de relação dos gêneros com a média do índice SWL	45
Figura 23 – Boxplot de relação dos gêneros com a média do índice SPIN	46
Figura 24 – Boxplot de relação das horas jogadas com a média do índices clínico GAD-7	46
Figura 25 – Boxplot de relação das horas jogadas com a média do índices clínico SWL .	46
Figura 26 – Perfis previstos pela rede baseado no índice GAD-7	47
Figura 27 – Perfis previstos pela rede baseado no índice SWL	48
Figura 28 – Perfis previstos pela rede baseado no índice SPIN	48
Figura 29 – Distribuição das previsões geradas pela rede para os perfis do índice GAD-7	49
Figura 30 – Distribuição das previsões geradas pela rede para os perfis do índice SWL .	50
Figura 31 – Distribuição das previsões geradas pela rede para os perfis do índice SPIN .	50

LISTA DE ABREVIATURAS E SIGLAS

CID-11	Classificação Internacional de Doenças
DNN	<i>Dense Neural Network</i>
FOMO	<i>Fear Of Missing Out</i>
GAD-7	<i>Generalized Anxiety Disorder 7-item</i>
HAM-A	<i>Hamilton Anxiety Rating Scale</i>
HAM-D	<i>Hamilton Depression Rating Scale</i>
IGDS9-SF	<i>Internet Gaming Disorder Scale-Short Form</i>
IQR	<i>Interquartile Range</i>
LIME	<i>Local Interpretable Model-Agnostic Explanations</i>
PHQ-9	<i>Patient Health Questionnaire-9</i>
PSQI	<i>Pittsburgh Sleep Quality Index</i>
RNA	Redes Neurais Artificiais
SHAP	<i>Shapley Additive Explanations</i>
SPIN	<i>Social Phobia Inventory</i>
SWL	<i>Satisfaction With Life</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Objetivo Geral	13
1.2	Objetivos Específicos	14
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Saúde mental segundo a OMS e o CID-11	16
2.2	Definição psicometrológica dos índices clínicos utilizados	16
2.3	O papel das emoções na dependência digital	17
2.4	Redes neurais artificiais: conceitos fundamentais	18
2.5	Redes neurais densas (<i>Dense Neural Network</i> (DNN))	19
2.6	Aplicação de redes neurais na análise de comportamentos em jogos online	19
3	TRABALHOS RELACIONADOS	21
3.1	Inteligência artificial e saúde mental	21
3.1.1	Aplicações de IA na psicologia	21
3.1.2	Modelos preditivos de ansiedade e aprendizado de máquina	22
3.2	Psicologia dos jogos e sistemas de recompensa	22
3.2.1	Fatores psicológicos em jogos digitais	22
3.2.2	Comportamentos ansiosos e o ambiente de jogo	23
3.2.3	Personalidade, motivação e recompensas	24
3.3	Aprendizado de máquina em dados comportamentais de jogadores	25
3.3.1	Estudos com dados de jogadores e plataformas digitais	25
3.3.2	Comportamento e engajamento em jogos digitais	25
3.4	Lacunas interdisciplinares	26
3.4.1	Lacunas na integração entre psicologia, jogos e IA	26
3.4.2	Limitações dos estudos existentes	27
3.5	Síntese dos trabalhos relacionados	27

4	METODOLOGIA	30
4.1	Conjunto de dados	30
4.2	Pré-processamento dos dados	31
4.2.1	Padronização semântica das respostas	31
4.2.2	Tratamento de valores ausentes e correlação	31
4.2.3	Codificação de variáveis	33
4.2.4	Normalização dos dados	33
4.2.5	Balanceamento dos índices	34
4.2.6	Análise de <i>outliers</i>	35
4.2.7	Distribuição de índices clínicos	35
4.3	Preparação para modelagem	38
4.4	Construção da rede neural	39
4.4.1	Arquitetura do modelo	39
4.4.2	Configuração do treinamento	41
4.5	Avaliação do modelo	41
5	RESULTADOS	43
5.1	Acurácia final de cada rede	43
5.2	Matriz de confusão e análise dos erros	43
5.3	Distribuição de características por índice clínico	45
5.4	Perfis típicos previstos	47
5.5	Distribuição das previsões por nível clínico	49
5.6	Ameaças à validade e limitações	50
6	CONCLUSÕES E TRABALHOS FUTUROS	52
6.1	Conclusões	52
6.2	Trabalhos futuros	52
	REFERÊNCIAS	54

1 INTRODUÇÃO

Segundo (Azevedo; Nascimento; Souza, 2014), toda tecnologia criada e aplicada coletivamente gerou impactos de diferentes magnitudes, promovendo transformações comportamentais, culturais e econômicas. A partir do século XIX, tais influências passaram a afetar também as interações humanas, impulsionadas pela mediação tecnológica que reduziu distâncias e acelerou significativamente o fluxo de informações. No final do século XX, a comunicação em massa, viabilizada principalmente pela internet, consolidou o nascimento de uma nova realidade, denominada popularmente de “mundo virtual”. Esse cenário alimentado pela facilidade de acesso e estímulo constante pode favorecer padrões de uso problemático da tecnologia, gerando comportamentos de ansiedade e impulsividade em relação ao seu uso.

Ademais, segundo (Barreto, 2025), a psicologia compreende o vício também como um mecanismo de fuga emocional, que busca a obtenção de prazer e a supressão da dor. Entre os fatores que contribuem para o desenvolvimento da dependência estão o prazer imediato, a evasão de problemas e influências genéticas. Partindo dessa perspectiva, o presente estudo direciona seu foco aos jogos, que, por meio de sua estrutura deliberadamente viciante e da promessa de recompensa imediata, são capazes de induzir comportamentos impulsivos que podem resultar em consequências devastadoras para os usuários, sustentando um sistema onde a prosperidade se ergue em detrimento da miséria de muitos.

As Redes Neurais Artificiais (RNA) se destacam como modelos computacionais capazes de identificar e representar padrões complexos a partir de dados de natureza diversa. Inspiradas no funcionamento biológico das redes neurais humanas, essas estruturas são compostas por unidades interconectadas que processam informações de forma paralela e adaptativa (Braga; Ludermir; Carvalho, 2007). Tal característica possibilita sua aplicação em contextos onde há múltiplas variáveis interdependentes, como no estudo de comportamentos digitais associados à ansiedade e à impulsividade.

1.1 Objetivo Geral

O objetivo geral deste trabalho é desenvolver modelos de redes neurais densas capazes de prever indicadores psicológicos validados, *Generalized Anxiety Disorder 7-item (GAD-7)*, *Social Phobia Inventory (SPIN)* e *Satisfaction With Life (SWL)*, a partir de padrões comportamentais extraídos de uso digital.

1.2 Objetivos Específicos

- Tratar, padronizar e preparar os dados brutos, garantindo consistência e qualidade da base.
- Aplicar técnicas de normalização e balanceamento para ajustar distribuições e evitar vieses.
- Construir e treinar modelos de redes neurais densas, definindo arquiteturas adequadas para cada índice clínico.
- Avaliar o desempenho preditivo dos modelos, considerando acurácia, estabilidade e erros por classe.
- Gerar perfis comportamentais previstos, identificando padrões típicos para cada nível clínico.

Embora existam estudos que investigam a relação entre tecnologia e saúde mental, observa-se uma ausência na integração de abordagens computacionais preditivas com indicadores clínicos validados, como o GAD-7, o SPIN e o SWL. Nesse sentido, este trabalho diferencia-se por propor o uso de redes neurais densas para modelar padrões comportamentais digitais capazes de prever níveis de ansiedade, fobia social e satisfação com a vida. Essa abordagem multidisciplinar contribui para aproximar a Psicologia e a Inteligência Artificial, oferecendo uma perspectiva quantitativa sobre fenômenos psicológicos.

Neste estudo, os indicadores clínicos selecionados representam dimensões centrais da saúde emocional. O GAD-7, o SPIN e o SWL são instrumentos amplamente utilizados em pesquisas psicológicas e permitem avaliar, respectivamente, sintomas de ansiedade generalizada, aspectos relacionados à ansiedade social e a percepção global de satisfação com a vida. A escolha desses três índices possibilita uma análise abrangente dos estados emocionais associados ao uso intenso de ambientes digitais, articulando indicadores negativos e positivos de saúde mental.

A literatura evidencia lacunas importantes tanto na psicologia digital quanto nas aplicações de IA. Estudos mostram que pesquisas em saúde mental ainda se apoiam majoritariamente em autorrelatos e dados clínicos tradicionais (Barreto, 2025), enquanto trabalhos sobre jogos digitais raramente integram métricas comportamentais objetivas (Sauter; Draschkow, 2022). Além disso, aplicações de aprendizado profundo em saúde mental apresentam limitada utilização de dados reais de plataformas digitais combinados a índices clínicos validados (Arrieta et al., 2020). Este estudo busca contribuir diretamente para esse espaço interdisciplinar.

A estrutura deste trabalho foi organizada para conduzir o leitor da contextualização teórica à análise dos resultados. Inicialmente, apresenta-se uma fundamentação que aborda dependência digital, emoções e comportamento em ambientes virtuais, bem como os princípios das redes

neurais densas utilizadas na modelagem. Em seguida, são discutidos estudos relacionados que tratam da aplicação de Inteligência Artificial na saúde mental. Posteriormente, descrevem-se os procedimentos metodológicos, incluindo coleta, tratamento e modelagem dos dados comportamentais. Na sequência, são apresentados e analisados os resultados do modelo preditivo, com suas implicações e limitações. Por fim, as considerações finais sintetizam as contribuições do estudo e indicam possíveis direções para pesquisas futuras.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Saúde mental segundo a OMS e o CID-11

A Organização Mundial da Saúde define saúde mental como um estado de bem-estar em que o indivíduo reconhece suas próprias capacidades, lida adequadamente com os estresses normais da vida, trabalha de maneira produtiva e contribui para sua comunidade (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2022). Essa definição adota uma perspectiva ampla e funcional, não vinculada a uma única abordagem teórica da psicologia, o que permite sua aplicação em diferentes contextos clínicos e de pesquisa.

Além disso, a Classificação Internacional de Doenças (CID-11) descreve formalmente categorias diagnósticas relacionadas aos construtos analisados neste estudo. Entre elas, destacam-se o Transtorno de Ansiedade Generalizada (código 6B00), associado aos sintomas avaliados pelo instrumento GAD-7 e o Transtorno de Ansiedade Social (código 6B04), diretamente relacionado ao SPIN. Esses referenciais oferecem um enquadramento clínico padronizado, permitindo situar manifestações emocionais relacionadas ao uso intensivo de tecnologias digitais dentro de critérios reconhecidos internacionalmente (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2019).

2.2 Definição psicometrológica dos índices clínicos utilizados

Os construtos psicológicos analisados neste estudo são definidos a partir dos próprios instrumentos que os mensuram, evitando a adoção de uma abordagem específica entre as diversas escolas teóricas da psicologia. A ansiedade generalizada, por exemplo, é operacionalizada segundo os critérios do instrumento GAD-7, desenvolvido por (Spitzer et al., 2025), que avalia sintomas como preocupação excessiva, tensão, inquietação e dificuldade de controle dos pensamentos ansiosos. De forma semelhante, o construto de ansiedade social é definido a partir do Social Phobia Inventory (SPIN), validado por (Connor et al., 2000), que mensura medo, evitação e desconforto intenso em situações de interação social. Já a satisfação com a vida é compreendida segundo o modelo proposto por (Diener et al., 2024), no qual o SWLS avalia a percepção subjetiva de bem-estar global e julgamento cognitivo da própria vida, sem implicações diagnósticas. Assim, cada índice utilizado possui fundamentos próprios, empiricamente validados e independentes de escolas teóricas específicas.

2.3 O papel das emoções na dependência digital

Segundo Azevedo, Nascimento e Souza (2014), a evolução tecnológica produziu impactos de distintas proporções, promovendo transformações significativas no comportamento, cultura e economia, dando destaque a globalização que iniciou no século XX. A facilidade de acesso à informação proporcionada pela internet modificou profundamente os padrões de sociabilidade e comportamento, já que o que antes demandava tempo e esforço, como comprar um jornal ou ir a um estabelecimento, passou a ser realizado de forma imediata por meio de uma simples pesquisa online.

Os autores descrevem que, por meio do uso patológico da internet, diversos comportamentos nocivos podem se manifestar, como a Dependência Cibersexual, caracterizada pelo vício em conteúdo adulto; a Dependência de Ciber Relacionamentos, na qual as interações sociais se limitam ao ambiente virtual; o vício em jogos online, sejam competitivos, cooperativos, de leilões, comércio ou apostas; a Sobrecarga de Informação, em que há navegação compulsiva e medo de ficar desatualizado, fenômeno conhecido como *Fear Of Missing Out* (FOMO); e a Dependência de Computadores ou outros dispositivos, marcada pelo uso obsessivo para consumo de conteúdo, jogos ou programação. É importante destacar que todos esses comportamentos derivam da busca por prazer e satisfação imediatos, desencadeando impulsos compulsivos que o indivíduo pode não conseguir controlar, resultando em sentimentos de culpa, desconforto e prejuízos em sua vida, assim como ocorre em outros tipos de vício.

A literatura psicológica reforça que a dependência digital exerce impacto direto sobre o funcionamento emocional dos indivíduos (Azevedo; Nascimento; Souza, 2014). Esse estudo discute como a exposição contínua a estímulos digitais, como ciclos de recompensa rápida, notificações frequentes e a necessidade de atualização constante, ativa mecanismos associados ao sistema de recompensa e aumenta a sensibilidade emocional a estados de alerta e ansiedade. Pesquisas recentes sobre comportamento digital e saúde emocional, realizadas em contextos de jogos e uso intensivo de internet (Gioia; Colella; Boursier, 2022; Zhao et al., 2021), indicam que ambientes digitais intensivos elevam irritabilidade, preocupação antecipatória e sensação de inadequação, especialmente quando há comparação social contínua ou pressão por desempenho, ampliando a vulnerabilidade emocional e reforçando o uso compulsivo como tentativa de regulação afetiva.

No contexto específico dos jogos digitais, investigações sobre envolvimento problemático com jogos e apostas online (Barreto, 2025) aponta que estruturas de progressão, ciclos de vitória

e derrota e recompensas imprevisíveis intensificam oscilações emocionais, gerando flutuações entre euforia, frustração e estresse competitivo. Essa perspectiva é reforçada por estudos baseados em dados de hábitos de jogadores (Sauter; Draschkow, 2022), que descrevem o uso de jogos como estratégia recorrente de alívio emocional. Embora essa prática possa oferecer conforto momentâneo, as evidências apontam que o alívio é frequentemente temporário (Gioia; Colella; Boursier, 2022; Zhao et al., 2021), fortalecendo a associação entre bem-estar e exposição prolongada ao ambiente digital. Esse padrão contribui para o surgimento de sintomas ansiosos e para a redução gradual da satisfação geral com a vida, reforçando a importância de investigar emoções no âmbito da dependência digital.

2.4 Redes neurais artificiais: conceitos fundamentais

As Redes Neurais Artificiais (RNA) constituem uma classe de modelos de aprendizado de máquina inspirados, de forma abstrata, no funcionamento do sistema nervoso biológico (Haykin, 2001; Braga; Ludermir; Carvalho, 2007). Em sua forma mais básica, uma RNA é composta por unidades de processamento chamadas neurônios artificiais, que recebem sinais de entrada, aplicam uma transformação matemática e produzem uma saída. Cada conexão entre neurônios possui um peso associado, que determina a influência de uma entrada na resposta final do modelo.

O comportamento de cada neurônio é descrito pela combinação linear dos sinais de entrada e pelos pesos sinápticos, seguida da aplicação de uma função de ativação. Essa função introduz não linearidade ao sistema, permitindo que a rede aprenda relações complexas e não triviais entre variáveis. Funções como *sigmoid*, *tanh* e *ReLU* são amplamente utilizadas devido às suas propriedades matemáticas e eficiência computacional (Dubey; Singh; Chaudhuri, 2021).

As RNAs organizam-se em camadas: uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. O processo de aprendizado ocorre pela atualização iterativa dos pesos, utilizando um algoritmo de otimização como o *backpropagation*, que ajusta os parâmetros com base no erro entre a saída prevista e a saída desejada.

Com a introdução de camadas ocultas adicionais e o aumento da capacidade computacional, surgiram arquiteturas mais profundas que permitiram capturar padrões de alta complexidade. É nesse contexto que se inserem as redes neurais densas (DNN), discutidas na próxima seção, as quais representam uma evolução natural das RNAs tradicionais e ampliam significativamente sua capacidade de modelagem.

2.5 Redes neurais densas (DNN)

As redes neurais densas, também chamadas de *fully connected networks* ou *deep neural networks*, representam uma evolução direta das redes neurais artificiais. Diferentemente de arquiteturas mais simples, as DNN organizam-se em múltiplas camadas ocultas, nas quais cada neurônio se conecta a todos os neurônios da camada anterior. Essa estrutura totalmente conectada permite que o modelo aprenda representações internas progressivamente abstratas, capturando padrões complexos e não lineares em conjuntos de dados de alta dimensionalidade.

O aprendizado em DNN ocorre por meio do ajuste dos pesos sinápticos com algoritmos baseados em gradiente, destacando-se o otimizador Adam (Kingma; Ba, 2014), amplamente utilizado por sua eficiência e estabilidade em problemas com muitos parâmetros. A profundidade do modelo e o uso de funções de ativação não lineares, como *ReLU* e suas variações (Dubey; Singh; Chaudhuri, 2021), ampliam a capacidade de generalização, permitindo que redes profundas representem relações sofisticadas entre variáveis mesmo quando tais relações não são evidentes de forma explícita.

Além disso, mecanismos de regularização, como *dropout* e normalização de camadas, têm sido empregados para reduzir sobreajuste e garantir um aprendizado mais robusto (Chollet, 2025). Esses recursos tornam as DNN especialmente adequadas para cenários em que os dados apresentam ruído, variabilidade comportamental ou distribuição assimétrica — características comuns em registros de uso digital.

A aplicação de DNN no contexto da saúde mental vem ganhando destaque, especialmente em estudos que buscam identificar padrões comportamentais associados a condições psicológicas. Pesquisas recentes demonstram o potencial de modelos profundos para reconhecer sinais de ansiedade, impulsividade e padrões emocionais em dados derivados de interações digitais e ambientais (Hong et al., 2023; Wang et al., 2023; Younas; Haq, 2025). Nesse sentido, as redes densas foram escolhidas neste trabalho por sua capacidade de modelar relações sutis entre variáveis comportamentais e os índices clínicos GAD-7, SPIN e SWL, possibilitando previsões estáveis e interpretáveis dentro do escopo da análise preditiva em saúde mental.

2.6 Aplicação de redes neurais na análise de comportamentos em jogos online

As redes neurais densas (DNN) têm sido amplamente utilizadas em pesquisas que buscam compreender comportamentos em ambientes virtuais, especialmente no contexto dos jogos online.

Esses modelos permitem identificar padrões complexos de interação, uma vez que são capazes de capturar relações não lineares entre variáveis comportamentais, sociais e emocionais. No campo da saúde mental, a aplicação de arquiteturas profundas possibilita explorar como fatores como engajamento digital, impulsividade, recompensas e tempo de uso podem se associar a estados psicológicos, incluindo ansiedade, fobia social e satisfação com a vida.

No contexto dos jogos online, abordagens baseadas em aprendizado profundo têm sido empregadas para analisar comportamentos associados ao uso problemático e seus impactos na saúde mental, incluindo modelos voltados à identificação de transtorno do jogo pela internet e à avaliação de efeitos psicológicos em jogadores (Hong et al., 2023; Wang et al., 2023; Younas; Haq, 2025). Esses modelos apoiam a compreensão teórica de como estímulos oferecidos pelos jogos, como ciclos de recompensa, progressão estruturada e feedback imediato, podem contribuir para o reforço de comportamentos impulsivos em indivíduos suscetíveis. Do ponto de vista computacional, redes neurais densas mostram-se especialmente adequadas para esse tipo de investigação, uma vez que foram concebidas para modelar relações não lineares complexas entre variáveis (Haykin, 2001; Braga; Ludermir; Carvalho, 2007), o que justifica sua adoção na modelagem preditiva desenvolvida neste estudo.

3 TRABALHOS RELACIONADOS

Este capítulo apresenta os principais trabalhos que serviram de fonte para fundamentar a pesquisa, além de demonstrar necessidade de aplicação do tema trabalhado, tratando também da correlação entre psicologia e tecnologia, demonstrando também, como o sistema de recompensas nos jogos está ligado à psicologia, além do uso do aprendizado de máquina para integração entre essas áreas. Ao final, é apresentada uma síntese que realça a contribuição e o diferencial deste estudo.

3.1 Inteligência artificial e saúde mental

3.1.1 Aplicações de IA na psicologia

Diversos estudos recentes têm demonstrado o potencial do aprendizado de máquina na detecção e análise de sintomas psicológicos. Guo et al. (2021) propuseram o uso de redes neurais profundas para reconhecer sinais de depressão a partir de expressões faciais, obtendo resultados promissores na diferenciação entre indivíduos saudáveis e aqueles com sintomas depressivos. De forma semelhante, Şipoş et al. (2024) aplicaram técnicas de aprendizado de máquina para analisar padrões de engajamento digital em adolescentes, identificando correlações significativas entre o comportamento online e sintomas de TDAH e ansiedade. Já TABARES et al. (2024) realizaram uma análise preditiva dos níveis de ansiedade em jovens com base em múltiplos indicadores comportamentais, evidenciando a eficácia de modelos de aprendizado profundo na previsão de estados emocionais. Esses resultados reforçam a relevância do uso de abordagens computacionais na psicologia contemporânea, indicando que as redes neurais e demais técnicas de aprendizado de máquina podem atuar como instrumentos de apoio à avaliação psicológica e à compreensão de fatores associados ao bem-estar mental.

Entre os modelos mais utilizados, destacam-se os índices clínicos padronizados como o GAD-7 (Spitzer et al., 2025), SPIN (Connor et al., 2000) e SWL (Diener et al., 2024). Como exemplo, Vliet (2022) investigaram a relação entre a satisfação com a vida e aspectos como fobia social e ansiedade, utilizando as escalas SWL, SPIN e GAD-7 para análise de correlação em jogadores. Os autores observaram que a ansiedade foi o fator mais relevante para a detecção de insatisfação com a vida, embora o modelo tenha apresentado erros em cerca de 21% dos casos, indicando que variáveis analisadas isoladamente possuem baixa capacidade preditiva.

Esses resultados reforçam a importância de considerar a correlação e a interação entre diferentes fatores no processo de classificação de perfis psicológicos.

3.1.2 Modelos preditivos de ansiedade e aprendizado de máquina

Estudos recentes têm explorado diferentes abordagens preditivas para sintomas emocionais, destacando a importância de técnicas como o balanceamento de classes (Ali et al., 2019) e a normalização dos dados (Amorim et al., 2022) para otimizar o desempenho dos modelos. O Avanço de arquiteturas de redes neurais profundas (Dubey; Singh; Chaudhuri, 2021) e o uso de funções de perda como a Softmax (Li et al., 2020) têm favorecido aplicações voltadas à detecção de padrões emocionais e clínicos.

Dentro desse contexto, alguns trabalhos se concentram especificamente na predição de transtornos de ansiedade, utilizando redes neurais artificiais para classificar indivíduos em diferentes níveis com base em questionários padronizados, como o GAD-7. Apesar dos resultados promissores, muitos desses estudos ainda utilizam dados clínicos controlados e carecem de variáveis comportamentais contextuais, o que limita a aplicabilidade prática em cenários digitais, como ambientes de jogos online.

3.2 Psicologia dos jogos e sistemas de recompensa

3.2.1 Fatores psicológicos em jogos digitais

Os jogos digitais utilizam uma variedade de mecanismos psicológicos para promover engajamento contínuo, baseando-se em princípios de condicionamento e reforço. Elementos como recompensas variáveis e desafios diários estimulam o comportamento de repetição e trazem a sensação de progresso constante e realização pessoal. Esses sistemas de recompensa, embora eficazes para retenção de jogadores, também podem intensificar sentimentos de ansiedade, especialmente quando associados à pressão por desempenho e à competição online (Barreto, 2025).

Além dos aspectos de reforço, os jogos frequentemente despertam respostas emocionais relacionadas à dopamina e à expectativa de conquista, fenômeno amplamente observado em estudos sobre dependência tecnológica e comportamento digital (Azevedo; Nascimento; Souza, 2014). A constante exposição a estímulos visuais e recompensas imprevisíveis reforça o desejo de permanecer conectado, levando o jogador a associar o retorno ao jogo a uma forma de alívio

emocional temporário.

O design psicológico dos jogos exerce influência direta sobre a experiência do jogador, modulando aspectos como: motivação, autoconfiança e percepção de controle. A busca por recompensas e reconhecimento social dentro do ambiente virtual tende a estimular comportamentos de comparação e frustração, especialmente em contextos de recompensas limitadas. Embora esses sistemas sejam fundamentais para a imersão e o entretenimento, também podem servir como gatilhos para desequilíbrios emocionais quando associados a perfis mais vulneráveis ou padrões de engajamento excessivo.

3.2.2 Comportamentos ansiosos e o ambiente de jogo

A relação entre ansiedade e jogos digitais envolve múltiplos fatores psicológicos e sociais. Ambientes competitivos e baseados em desempenho tendem a intensificar sentimentos de tensão e insegurança, especialmente quando há comparações constantes entre jogadores ou recompensas atreladas à performance. Esse tipo de estrutura pode reforçar o medo de falhar ou de ser superado, estimulando comportamentos de vigilância e autocobrança. Em jogos com rankings ou sistemas de pontuação pública, o impacto emocional é amplificado pela pressão social e pela necessidade de reconhecimento.

Em espaços de interação online, especialmente em jogos competitivos, a exposição contínua à avaliação de outros jogadores, comparações de desempenho e interações hostis pode intensificar sintomas de ansiedade social e mal-estar emocional, sobretudo em indivíduos vulneráveis (Gioia; Colella; Boursier, 2022; Zhao et al., 2021). Dados comportamentais provenientes de plataformas de jogo evidenciam essas dinâmicas: o *Online Gaming Anxiety Dataset* (Divyansh22, 2021), apresenta correlações entre hábitos de jogo, isolamento digital e sintomas de ansiedade, reforçando a relevância de analisar o comportamento em ambientes virtuais sob uma perspectiva psicológica.

Além disso, o próprio ato de jogar pode funcionar como uma estratégia de regulação emocional. Muitos jogadores recorrem aos jogos como uma forma de aliviar estresse, solidão ou frustração, o que pode gerar uma dependência emocional do ambiente digital. Embora o jogo ofereça um espaço temporário de controle e previsibilidade, a repetição desse padrão tende a reforçar o ciclo de ansiedade, especialmente quando o jogador percebe o retorno constante ao jogo como a única forma eficaz de lidar com desconfortos emocionais.

3.2.3 Personalidade, motivação e recompensas

A compreensão da motivação e dos traços de personalidade dos jogadores é fundamental para interpretar o modo como cada indivíduo reage aos sistemas de recompensa digital. O modelo proposto por (Bartle, 1996) descreve diferentes perfis de jogador com base em motivações relacionadas à competição, exploração e socialização, oferecendo uma perspectiva útil para compreender como cada tipo de engajamento pode influenciar emoções e comportamentos durante o jogo. Essa tipologia foi posteriormente expandida de forma empírica por (Yee, 2006), que identificou três dimensões principais de motivação em jogos online: *Achievement*, *Social* e *Immersion*. A dimensão *Achievement* reúne motivações ligadas ao progresso, à competição e ao domínio das mecânicas do jogo; *Social* envolve aspectos de interação, cooperação e formação de relacionamentos; e *Immersion* abrange fatores relacionados à exploração, à personalização e ao escapismo. Esse modelo amplia a estrutura teórica de Bartle ao oferecer uma base empírica que explica como diferentes formas de motivação influenciam o envolvimento emocional e o comportamento dos jogadores em ambientes virtuais. Estudos complementares também relacionam traços de personalidade como impulsividade e busca por estímulo, à predisposição para comportamentos ansiosos ou compulsivos em contextos de jogo (Azevedo; Nascimento; Souza, 2014; Barreto, 2025).

Outros autores abordam a influência dos fatores motivacionais sobre o bem-estar digital e o engajamento emocional. (Vliet, 2022) associa a satisfação de vida e a estabilidade emocional ao equilíbrio entre metas pessoais e experiência de lazer, sugerindo que o jogo pode tanto promover bem-estar quanto intensificar vulnerabilidades psicológicas, dependendo do perfil motivacional do jogador. De modo semelhante, (Leung; Lee, 2005) investigaram como a busca por conexões sociais no meio digital pode gerar dependência emocional e reforçar sentimentos de isolamento, fenômeno que encontra paralelo nos sistemas de recompensa e validação presentes nos jogos online.

O autor Barreto (2025) observa que a exposição prolongada a recompensas condicionais e a interações competitivas pode gerar desequilíbrios emocionais associados à comparação social e à autocrítica. O autor destaca que o prazer obtido em jogos de progressão contínua tende a se confundir com o alívio de tensão, tornando o reforço psicológico mais intenso e menos consciente.

3.3 Aprendizado de máquina em dados comportamentais de jogadores

3.3.1 Estudos com dados de jogadores e plataformas digitais

O uso de aprendizado de máquina para analisar dados de jogadores tem se expandido à medida que as plataformas digitais disponibilizam métricas detalhadas sobre comportamento, engajamento e tempo de jogo. Trabalhos recentes demonstram que informações como gênero de jogo, frequência de login, tempo de sessão e interações sociais podem ser utilizadas para prever padrões de retenção e perfis de uso (Sauter; Draschkow, 2022). Além disso, estudos aplicados têm mostrado que modelos de aprendizado profundo conseguem identificar padrões emocionais e comportamentais em jogadores, especialmente em contextos de uso problemático, como demonstrado por Hong et al. (2023) e Wang et al. (2023), que empregam redes neurais para detectar indicadores de transtorno de jogos digitais.

Pesquisas também têm utilizado dados coletados em comunidades e redes sociais associadas a jogos, o que permite compreender a influência de fatores sociais e contextuais sobre o comportamento digital. (Leung; Lee, 2005) discutem que a participação em ambientes virtuais pode estar associada à busca por pertencimento, mas também à intensificação de sintomas ansiosos quando há exposição contínua à comparação social. De modo semelhante, (Vliet, 2022) argumenta que o equilíbrio entre lazer e desempenho influencia diretamente o bem-estar emocional em contextos digitais, o que reforça a importância de interpretar os dados comportamentais sob uma perspectiva psicológica.

Modelos computacionais aplicados a essas bases têm alcançado resultados expressivos na identificação de padrões complexos de engajamento e evasão, sugerindo que o comportamento digital pode refletir características emocionais e motivacionais dos jogadores. Pesquisas mais recentes reforçam esse potencial ao empregar aprendizado de máquina para estimar estados emocionais e efeitos psicológicos relacionados ao ato de jogar (Younas; Haq, 2025). Assim, o aprendizado de máquina emerge como uma ferramenta relevante não apenas para prever retenção ou consumo, mas também para compreender dimensões psicológicas associadas à interação com jogos online.

3.3.2 Comportamento e engajamento em jogos digitais

A análise do engajamento em jogos digitais envolve tanto fatores cognitivos quanto emocionais, sendo influenciada pela estrutura do jogo, pelos sistemas de recompensa e pelo

perfil motivacional de cada jogador. Estudos baseados em dados de plataformas, como o *Gaming Habits Dataset* (Sauter; Draschkow, 2022), mostram que o tempo de exposição e o tipo de jogo estão diretamente relacionados à satisfação e ao cansaço emocional, sugerindo que padrões de comportamento digital podem servir como indicadores indiretos de bem-estar ou desgaste psicológico.

Além das métricas quantitativas, pesquisas apontam que o engajamento é afetado por elementos subjetivos, como senso de progresso, cooperação e imersão narrativa. (Vliet, 2022) destaca que o equilíbrio entre prazer e controle é um fator essencial para que a experiência de jogo seja positiva, evitando a transição de lazer para compulsão. Em contrapartida, (Barreto, 2025) ressalta que a sobreposição entre desafio e recompensa pode gerar uma percepção distorcida de satisfação, na qual o jogador busca o jogo não pelo entretenimento, mas como mecanismo de regulação emocional.

Essas evidências indicam que o comportamento em jogos digitais é multifatorial, resultante da interação entre características individuais, design do jogo e contexto social. Dessa forma, compreender o engajamento requer uma abordagem que una dados comportamentais e interpretações psicológicas, possibilitando a identificação de padrões que vão além das métricas tradicionais de desempenho.

3.4 Lacunas interdisciplinares

3.4.1 Lacunas na integração entre psicologia, jogos e IA

Apesar do avanço das pesquisas nas áreas de psicologia, jogos digitais e inteligência artificial, ainda há uma ausência significativa de estudos que integrem essas três dimensões de forma estruturada. Trabalhos sobre saúde mental frequentemente limitam-se a contextos clínicos e autorrelatos, enquanto pesquisas em jogos tendem a priorizar aspectos técnicos ou de design. Do mesmo modo, aplicações de aprendizado de máquina em psicologia costumam concentrar-se em diagnósticos clínicos convencionais, desconsiderando variáveis comportamentais extraídas de ambientes digitais.

Autores como (Azevedo; Nascimento; Souza, 2014) e (Vliet, 2022) destacam a necessidade de abordagens interdisciplinares que combinem dados objetivos de comportamento digital com indicadores emocionais, de modo a ampliar o entendimento sobre como a interação com sistemas de recompensa influencia o bem-estar. No entanto, essa integração ainda é de nível

muito inicial, o que reforça a relevância de modelos que empreguem redes neurais para identificar correlações entre comportamento, motivação e estados psicológicos.

Nesse contexto, o presente trabalho propõe preencher parte dessa lacuna ao utilizar técnicas de aprendizado profundo para prever índices clínicos de ansiedade com base em dados comportamentais de jogadores, contribuindo para o avanço de pesquisas que unem psicologia e tecnologia.

3.4.2 Limitações dos estudos existentes

Instrumentos amplamente utilizados, como o GAD-7 (Spitzer et al., 2025), o SPIN (Connor et al., 2000) e o SWL (Diener et al., 2024), baseiam-se inteiramente na autoavaliação dos participantes, não incorporando informações comportamentais objetivas.

Outros trabalhos também restringem suas análises a bases de dados específicas, o que pode limitar a generalização dos resultados em determinados contextos. Por exemplo, Sauter e Draschkow (2022) utilizam informações provenientes de um único conjunto de jogadores para investigar padrões de engajamento. Esse tipo de limitação é recorrente na área, refletindo a disponibilidade de dados estruturados sobre comportamento digital e não necessariamente uma fragilidade metodológica isolada, mas uma condição comum em estudos que trabalham com bases reais provenientes de plataformas de jogo.

Além disso, revisões recentes indicam que ainda há uma escassez de estudos que integrem abordagens explicáveis de inteligência artificial em áreas sensíveis, (Arrieta et al., 2020), sobretudo em aplicações que envolvem comportamento humano e saúde mental. A literatura destaca que, embora modelos preditivos avancem rapidamente, sua interpretabilidade permanece limitada, gerando desafios na contextualização gerando desafios na contextualização psicológica das previsões (Molnar, 2022).

3.5 Síntese dos trabalhos relacionados

A literatura analisada revela avanços importantes em três frentes principais: o uso de inteligência artificial na avaliação psicológica, a compreensão dos mecanismos emocionais associados aos jogos digitais e a aplicação de aprendizado de máquina sobre dados comportamentais de jogadores.

Na área de saúde mental, estudos recentes têm demonstrado que modelos de aprendizado

profundo podem prever sintomas como depressão, ansiedade e TDAH a partir de sinais comportamentais ou expressões faciais (Guo et al., 2021; Şipoş et al., 2024; TABARES et al., 2024). Esses trabalhos evidenciam o potencial das técnicas de IA como instrumentos auxiliares em processos clínicos, reforçando a relevância de índices padronizados como GAD-7, SPIN e SWL na análise de perfis emocionais (Spitzer et al., 2025; Connor et al., 2000; Diener et al., 2024). Estudos como o de Vliet (2022) demonstram ainda que variáveis psicológicas interagem de forma complexa e que análises isoladas apresentam limitações preditivas, reforçando a necessidade de abordagens multivariadas.

No campo da psicologia dos jogos, pesquisas apontam que os sistemas de recompensa, a competição e as dinâmicas sociais dos ambientes virtuais exercem influência significativa sobre emoções como ansiedade, frustração e motivação (Barreto, 2025; Azevedo; Nascimento; Souza, 2014). Esses elementos afetam tanto o engajamento quanto o bem-estar emocional, especialmente em contextos de exposição contínua a comparações, pressão por desempenho e interações hostis, como demonstram Gioia, Colella e Boursier (2022) e Zhao et al. (2021). Perfis motivacionais e traços de personalidade também moldam essas respostas, conforme discutido por (Bartle, 1996; Yee, 2006), ampliando o entendimento sobre diferenças individuais na forma como jogadores se relacionam com sistemas digitais.

Por fim, estudos envolvendo aprendizado de máquina aplicado diretamente a dados comportamentais de jogadores mostram que métricas como tempo de jogo, frequência, padrões de login e interações sociais podem revelar indícios de uso problemático, estados emocionais e até sinais de sofrimento psicológico (Sauter; Draschkow, 2022; Hong et al., 2023; Wang et al., 2023). Trabalhos mais recentes sugerem que modelos preditivos podem estimar indicadores emocionais a partir de variáveis comportamentais em plataformas de jogo (Younas; Haq, 2025), embora ainda haja limitações quanto à disponibilidade de bases amplas e quanto à presença de métodos explicáveis de IA que permitam interpretar de forma transparente os mecanismos subjacentes às previsões (Arrieta et al., 2020; Molnar, 2022).

Em conjunto, os estudos revisados mostram que, embora existam contribuições significativas em cada uma das três áreas: psicologia, jogos digitais e aprendizado de máquina, poucos trabalhos integram esses domínios de maneira conjunta. Assim, permanece uma lacuna importante na literatura: a falta de modelos que utilizem dados comportamentais reais de jogadores para prever indicadores clínicos de saúde mental de forma interpretável e aplicável ao contexto digital.

O presente trabalho busca preencher parte dessa lacuna ao empregar redes neurais densas para prever níveis clínicos de ansiedade a partir de variáveis comportamentais, oferecendo uma abordagem interdisciplinar que combina técnicas de aprendizado profundo, fundamentos psicológicos e análise de dados comportamentais provenientes de jogadores. Essa integração permite avançar na compreensão de como padrões de interação digital se relacionam com estados emocionais, destacando o potencial da inteligência artificial como ferramenta de apoio em estudos sobre bem-estar em ambientes virtuais.

4 METODOLOGIA

Este estudo empregou uma abordagem quantitativa com suporte computacional, utilizando técnicas de aprendizado de máquina e procedimentos de ciência de dados para analisar e modelar padrões comportamentais relacionados à ansiedade e impulsividade em ambientes de jogos online.

A linguagem de programação escolhida foi Python, amplamente utilizada por sua simplicidade e pela variedade de bibliotecas voltadas ao processamento, análise e modelagem de dados, dentre as quais foram selecionadas: *pandas* (manipulação dos dados), *numpy* (operações matemáticas e vetorização para performance e cálculos estatísticos), *matplotlib* e *seaborn* (geração dos gráficos e visualização dos dados), *scikit-learn* (para pré-processamento, codificação, escalonamento, divisão dos dados e avaliação dos modelos feitos), *TensorFlow/Keras* (construção, treinamento e validação da rede neural).

O processo metodológico dividiu-se em três etapas principais: coleta e compreensão dos dados, pré-processamento e preparação dos dados para modelagem preditiva.

4.1 Conjunto de dados

Os dados utilizados foram obtidos a partir do *dataset Online Gaming Anxiety Data*¹, disponibilizado publicamente na plataforma Kaggle (Divyansh22, 2021). Este conjunto foi originalmente estruturado por Marian Sauter e Dejan Draschkow, tratando-se de uma pesquisa aplicada ao redor do mundo (Sauter; Draschkow, 2022). O motivo da escolha é que a base, além de obter diversos perfis de jogadores, também conta com informações relevantes, onde a estruturação dos questionários que compuseram a pesquisa, deu-se principalmente por meio de escalas clínicas indicadores de ansiedade (medidos pela escala GAD-7), fobia social (SPIN) e satisfação com a vida (SWL).

Os questionários integram um estudo psicológico hospedado no site *Center for Open Science*², que posteriormente foram disponibilizados pelos autores e a base final foi montada por meio de plataformas e comunidades virtuais, como: Reddit, TeamLiquid, CrowdFlower, além de outros fóruns relacionados a jogos.

O arquivo disponibilizado é composto por 55 colunas e contém respostas autodeclaradas de jogadores acerca de sua experiência com jogos online, motivação para jogar, frequência

¹ Disponível em: <https://www.kaggle.com/datasets/divyansh22/online-gaming-anxiety-data>

² Disponível em: <https://osf.io/vnbxk/overview>

de jogo semanal, tipo de interação durante o jogo, os 3 índices mencionados anteriormente (GAD-7, SWL, SPIN), além de características sociodemográficas que compõem o perfil de cada participante, sendo capaz de associá-los às tendências psicológicas.

4.2 Pré-processamento dos dados

4.2.1 Padronização semântica das respostas

Inicialmente, os dados foram inspecionados quanto à consistência e completude. As colunas categóricas foram padronizadas manualmente com o objetivo de reduzir ruídos oriundos de variações semânticas ou ortográficas. Por exemplo, diferentes variações da liga (presente na coluna "league") "Gold"(como "Gold II"ou "Gold 5") foram unificadas sob um único rótulo representativo ("Gold"), a fim de manter a coerência da análise.

Ademais, os motivos para os usuários jogarem (coluna "whyplay") também foram normalizados, onde motivos fora do escopo principal foram categorizados como “outros”, o mesmo acontecendo na coluna "playstyle", onde os jogadores descrevem se jogam jogos que envolvem outros jogadores ou não, e se costumam jogar sozinhos, com amigos, ou com grupos online.

4.2.2 Tratamento de valores ausentes e correlação

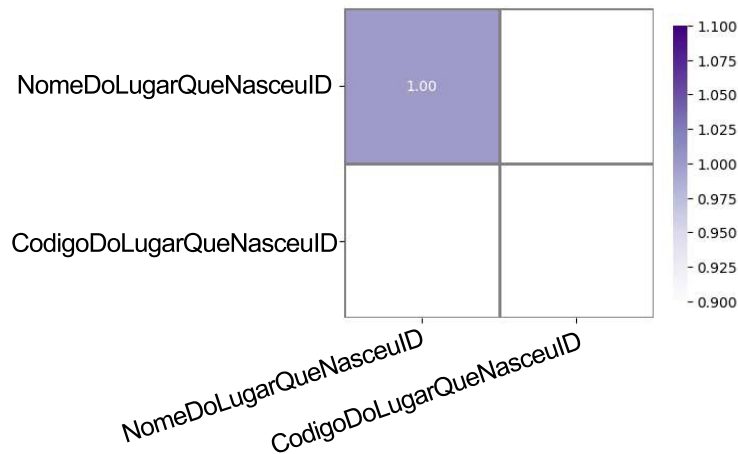
Durante a etapa de inspeção inicial, verificou-se a ocorrência de valores ausentes em determinadas colunas, demandando estratégias específicas de tratamento. Por exemplo, a coluna “highestleague”, que representa o nível mais alto da liga alcançado pelo jogador, que destaca-se por estar completamente vazia, foi excluída do conjunto de dados dado o seu caráter irrelevante para as etapas posteriores. Também nas colunas relacionadas ao índice SPIN, que possuem alguns valores ausentes. Tal análise foi essencial para tomada de decisões posteriores no descarte e imputação seletiva de valores e variáveis.

Durante o pré-processamento, identificou-se um problema de duplicidade e inconsistência nas colunas “Birthplace” e “Residence”, as quais representavam respectivamente, o país de nascimento e o país de residência dos participantes. Ambas as versões possuíam outras colunas correspondentes mas em formato padronizado ISO3, mas não estavam sincronizadas corretamente.

Visando solucionar essa inconsistência, foi criado um dicionário de correspondência entre

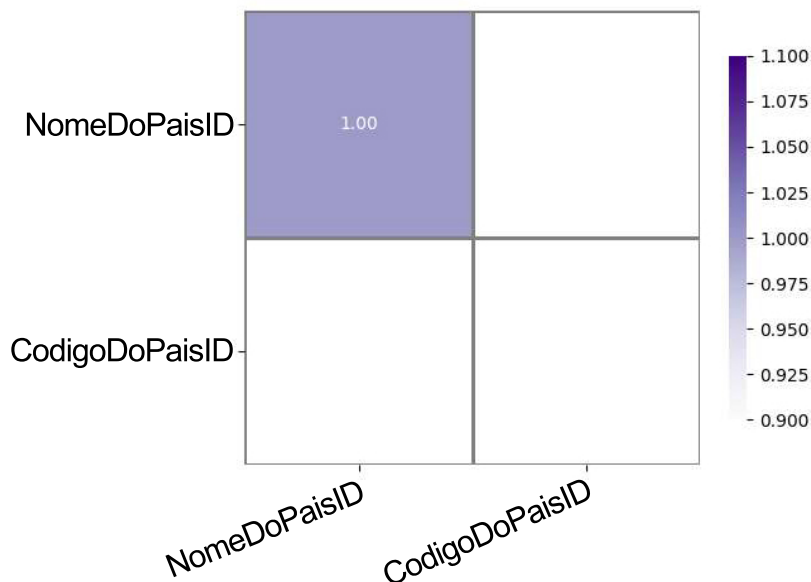
o nome dos países e seus respectivos códigos em ISO3, com base nas combinações não nulas e únicas observadas no *dataset*. A partir deste mapeamento, procedeu-se à correção dos valores inconsistentes ou ausentes, substituindo-os pela forma padronizada. Em seguida, foi realizado um mapeamento numérico com atribuição de valor único para cada país, permitindo assim criar um *heatmap* de correlação (Figura 1 e Figura 2), para demonstrar o grau de similaridade entre os valores nominais e os valores codificados em ISO3.

Figura 1 – *Heatmap* da coluna *Birthplace* após organização



Fonte: Autoria Própria (2025).

Figura 2 – *Heatmap* da coluna *Residence* após organização



Fonte: Autoria Própria (2025).

4.2.3 Codificação de variáveis

Para colunas com valores clínicos, evitou-se substituir dados ausentes por medidas estatísticas, como a média ou a moda, conforme a natureza dos dados, uma vez que aplicar qualquer técnica de substituição resultaria no viés da média desses valores para os usuários.

Com o objetivo de preservar os dados ausentes presentes nas colunas relacionadas ao índice SPIN, como a escala varia de 0 a 68, todos os valores faltantes identificados foram substituídos por 0. Essa abordagem preserva o significado clínico da escala, evitando distorções por médias artificiais.

Embora a imputação pela média não elimine viés e possa reduzir parcialmente a variabilidade natural dos dados, essa estratégia foi adotada neste estudo por manter o formato geral da distribuição original das variáveis numéricas e evitar distorções mais severas que ocorreriam com técnicas mais agressivas de substituição. Como a proporção de valores ausentes era pequena, o impacto estatístico dessa escolha torna-se limitado, permitindo preservar a estrutura global dos dados sem introduzir padrões artificiais relevantes.

Para colunas categóricas, empregou-se a estratégia de preenchimento probabilístico, substituindo os valores ausentes por categorias sorteadas aleatoriamente, conforme a natureza dos dados válidos. Inicialmente, os valores foram codificados utilizando o `LabelEncoder` e tratados como zero temporariamente. Posteriormente, realizou-se uma verificação para assegurar que os zeros não correspondiam a uma categoria válida, o que permitiu sua substituição adequada como valores ausentes da coluna.

A escolha pela codificação das variáveis categóricas por meio do algoritmo `LabelEncoder`, deve-se à compatibilidade com o modelo proposto DNN, uma vez que esse tipo de arquitetura opera de forma mais eficiente com entradas numéricas inteiras ao invés de valores textuais.

4.2.4 Normalização dos dados

Após o tratamento dos dados e a codificação numérica dos valores categóricos para numéricos em toda base, iniciou-se o processo de normalização das variáveis contínuas por meio do algoritmo `StandardScaler`. Essa técnica realiza a padronização dos dados, ajustando os valores para uma distribuição com média igual a zero e desvio padrão igual a um. Esse método, conhecido como *z-score* (viabiliza a identificação da distância que o valor está da média), permite

a melhora na convergência do modelo e impede que variáveis com escalas distintas dominem o gradiente durante o treinamento.

A escolha do `StandardScaler` deu-se pela eficácia comprovada em arquiteturas de Redes Neurais Densas (DNN), nas quais a padronização baseada em *z-score* tende a produzir resultados superiores em relação a técnicas como `MinMaxScaler`, principalmente em cenários com presença de *outliers* ou distribuições não uniformes (Amorim et al., 2022).

Durante o processo de normalização, decidiu-se não aplicar a transformação nas variáveis clínicas (GAD-7, SPIN, SWL), a fim de preservar a integridade das escalas psicométricas originais. Esses índices foram utilizados posteriormente na interpretação clínica dos resultados. Qualquer alteração na escala poderia comprometer a validade semântica dos níveis mensurados.

4.2.5 Balanceamento dos índices

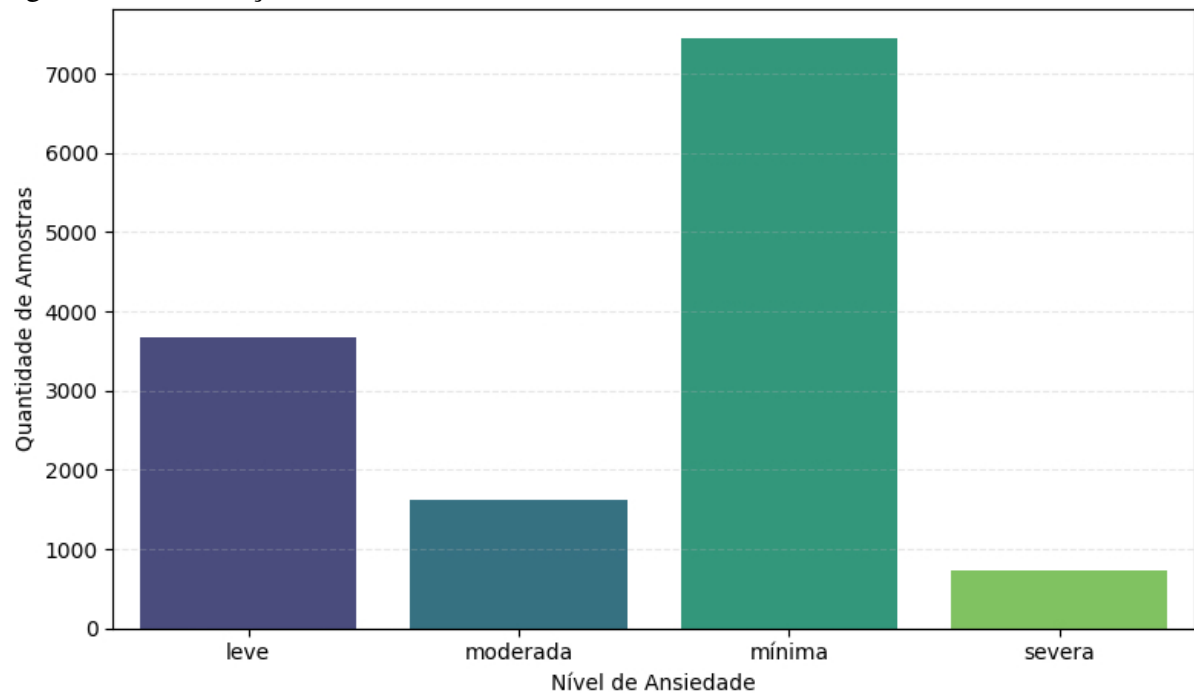
Para evitar que o modelo treinado seja tendencioso a um perfil em específico, o balanceamento dos dados foi feito, diminuindo a repetição de um resultado e forçando o modelo a analisar todos os casos existentes.

Primeiro, os valores de cada classificação foram analisados para criar uma nova coluna, que corresponde à interpretação dos seus resultados, por exemplo: no GAD-7, valores menores ou iguais a 4 resultam em chance mínima, entre 5 e 9 resultam em ansiedade leve, entre 10 e 14, moderada, e acima de 15, severa. Após a criação da coluna com essa interpretação, a distribuição dos níveis é apresentada na Figura 3.

Para compensar a desvantagem numérica dos níveis mais baixos, o método de balanceamento *oversampling* (Ali et al., 2019) foi escolhido, pois, com ele, os dados originais não se perdem, e sim, novos casos de perfis são criados para reforçar o aprendizado dos níveis com menos exemplos, mantendo a viabilidade da base original.

O tipo de *oversampling* utilizado foi o tipo aleatório, onde, segundo (Ali et al., 2019), perfis já existentes são duplicados até que todos os perfis tenham a mesma quantidade, evitando a criação de perfis fantasiosos que não correspondem à realidade da base de dados, valendo salientar que, a técnica só foi aplicada apenas nas colunas contendo os rótulos dos graus de cada índice, mantendo as colunas originais da base intactas. Após a utilização dessa técnica no GAD-7, SPIN e SWL, os perfis utilizados para treinamento estão apresentados na Figura 4.

Figura 3 – Distribuição inicial dos níveis de ansiedade



Fonte: Autoria Própria (2025).

4.2.6 Análise de *outliers*

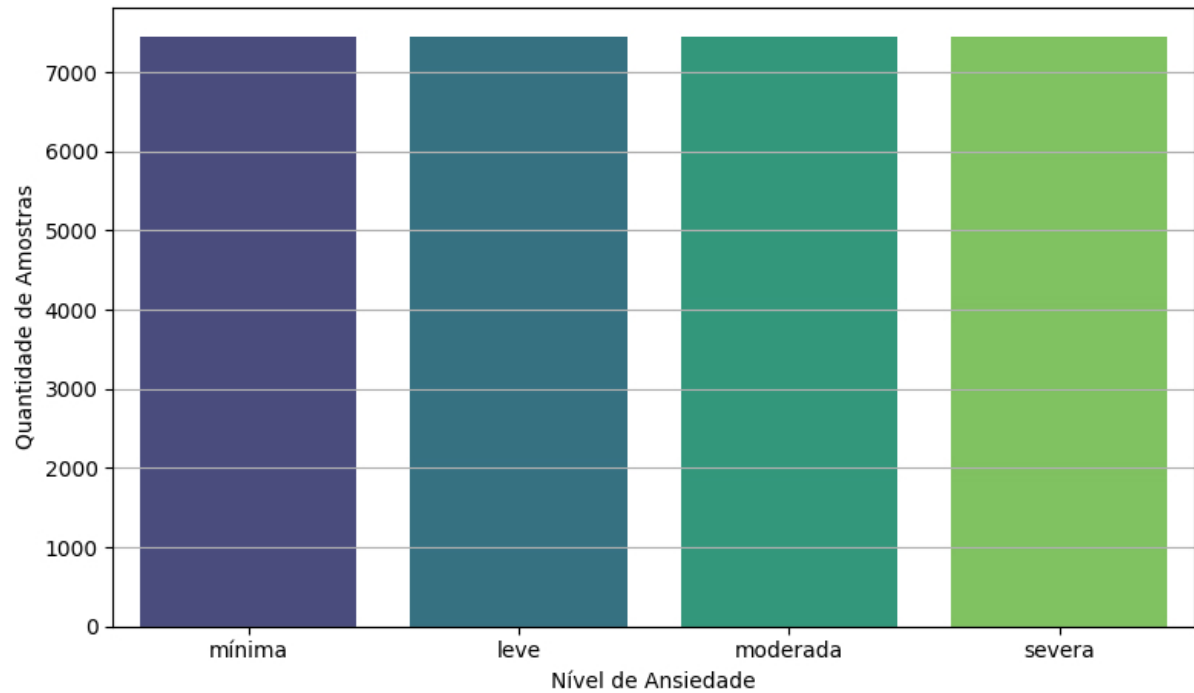
Foi realizada uma análise estatística para detectar e tratar valores extremos (*outliers*) nas colunas “streams”, “Age” e “Hours”. Utilizou-se a técnica baseada no intervalo interquartil *Interquartile Range* (IQR), onde valores que fossem inferiores a $Q1 - 1.5 \times IQR$ ou superiores a $Q3 + 1.5 \times IQR$, foram considerados anômalos. Com o objetivo de suavizar possíveis distorções, efetuou-se o emprego da mediana da respectiva variável, sem afetar a tendência central dos dados.

A Figura 5 ilustra a distribuição da coluna “streams” antes do tratamento de valores extremos baseando-se no IQR e a Figura 6 mostra como ficou a coluna após o tratamento. Percebe-se que, a substituição pela mediana suavizou a presença de valores desviantes, mantendo a estrutura geral da coluna. O mesmo procedimento foi adotado nas colunas “Age” (Figura 7 e Figura 8) e “Hours” (Figura 9 e Figura 10), respeitando suas respectivas distribuições.

4.2.7 Distribuição de índices clínicos

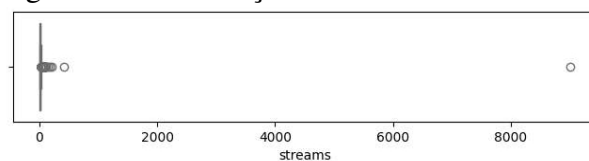
Com o objetivo de compreender a natureza dos dados clínicos utilizados neste estudo, foram gerados histogramas para as variáveis de média dos índices clínicos (“GAD-7”, “SPIN”,

Figura 4 – Distribuição após o balanceamento dos valores dos níveis de ansiedade



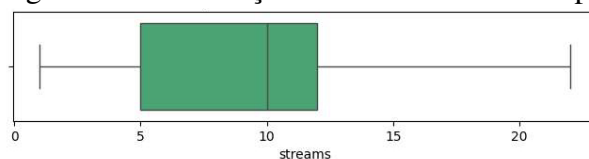
Fonte: Autoria Própria (2025).

Figura 5 – Distribuição da variável *stream* antes do tratamento



Fonte: Autoria Própria (2025).

Figura 6 – Distribuição da variável *stream* depois do tratamento



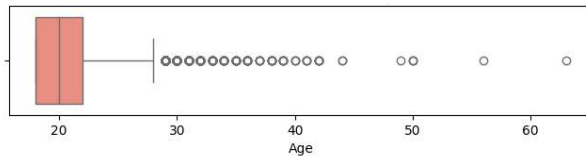
Fonte: Autoria Própria (2025).

“SWL”).

Os gráficos de distribuição apresentam a frequência observada para cada valor ou faixa de valores de um determinado índice clínico. Esse tipo de visualização permite compreender a forma da distribuição (assimetria, concentração, extremos) e identificar possíveis desequilíbrios entre categorias, o que é fundamental para avaliar desafios relacionados ao aprendizado do modelo.

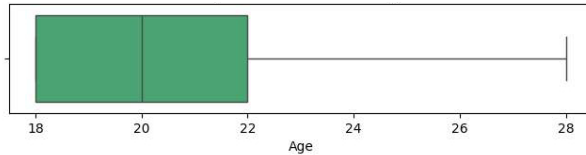
As distribuições observadas revelam escores baixos nas escalas de “GAD-7” e “SPIN” (Figura 11 e Figura 12), indicando sintomas mínimos a leves. A escala “SWL” (Figura 13),

Figura 7 – Distribuição da variável *Age* antes do tratamento



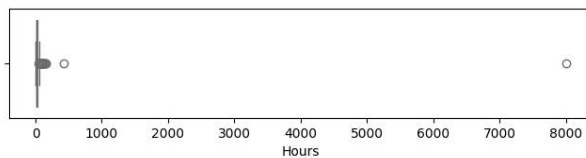
Fonte: Autoria Própria (2025).

Figura 8 – Distribuição da variável *Age* depois do tratamento



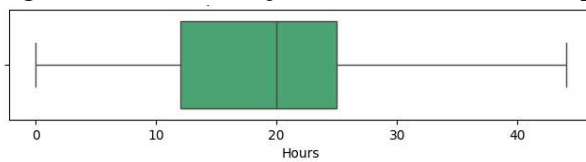
Fonte: Autoria Própria (2025).

Figura 9 – Distribuição da variável *Hours* antes do tratamento



Fonte: Autoria Própria (2025).

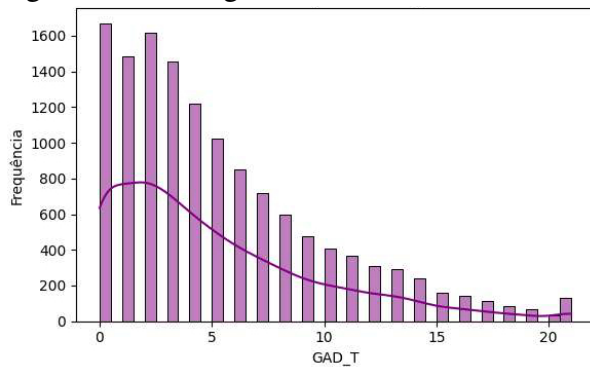
Figura 10 – Distribuição da variável *Hours* depois do tratamento



Fonte: Autoria Própria (2025).

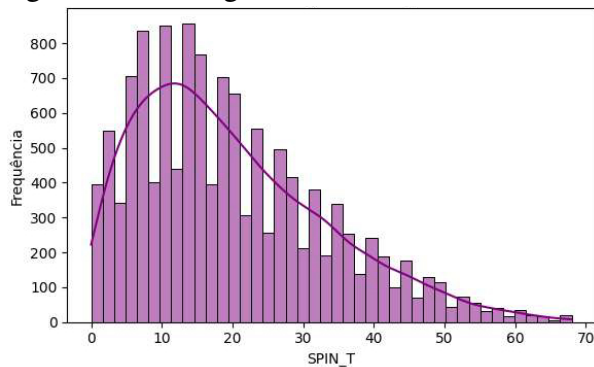
por sua vez, mostrou ter uma distribuição mais simétrica e concentrada em torno de escores intermediários, sugerindo níveis moderados de satisfação com a vida.

Figura 11 – Histograma do índice clínico GAD-7



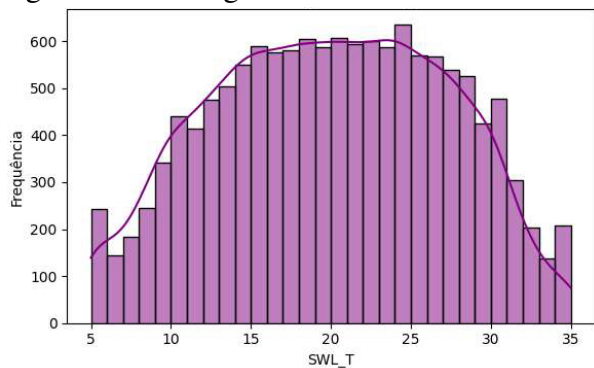
Fonte: Autoria Própria (2025).

Figura 12 – Histograma do índice clínico SPIN



Fonte: Autoria Própria (2025).

Figura 13 – Histograma do índice clínico SWL



Fonte: Autoria Própria (2025).

Tais padrões são esperados em coleta aberta e não clínica, ainda assim, a assimetria dos dados foi considerada nas etapas posteriores, aplicando-se uma estratégia de balanceamento para viabilizar um modelo capaz de reconhecer e prever com precisão todos os níveis clínicos, inclusive os menos representados.

Essa análise exploratória é essencial para garantir a transparência e orientar a interpretação dos resultados preditivos, sobretudo pelo fato de que tais variáveis não foram normalizadas.

4.3 Preparação para modelagem

A base foi dividida entre conjunto de treinamento e teste utilizando a função `train_test_split`, com proporção de 80% para treino e 20% para teste. Essa etapa garantiu que o modelo pudesse ser treinado com robustez e posteriormente avaliado de forma imparcial. As colunas alvo serão as médias das variáveis clínicas (“GAD-7”, “SPIN”, “SWL”), visto que o objetivo da rede é prever essas variáveis a partir das médias das respostas dos participantes:

- **GAD-7:** Os escores foram classificados por meio das suas quatro categorias: mínima, leve, moderada e severa. Ela foi desenvolvida como um instrumento para rastreamento e avaliação da gravidade da ansiedade generalizada, criada com base nos critérios do DSM-IV, com o objetivo de facilitar o diagnóstico clínico em contextos ambulatoriais.
- **SPIN:** É destinada a medir sintomas da ansiedade social de forma confiável e eficiente, avaliando sintomas fisiológicos, medo e comportamentos de esquiva relacionados à fobia social. Podendo ser usada como rastreio ou avaliação do tratamento.
- **SWL:** Tem como objetivo capturar a avaliação cognitiva que o indivíduo faz sobre sua própria vida como um todo, ignorando emoções específicas.

A escolha dessas variáveis buscou uma forma prática de treinar o modelo, visto que o resultado de cada uma foi autodeclarado por cada participante, fazendo com que seja mais fácil de observar a relação dos resultados com os atributos que compõem o perfil de cada jogador.

Para compor os perfis dos jogadores, as variáveis escolhidas foram: “game” (jogo mais jogado), “plataform” (plataforma onde costuma jogar, como computador ou console), “hours” (horas jogadas por semana), “whyplay” (motivação principal para jogar / presença de monetização), “league” (classificação, ou ranking, dentro do jogo), “streams” (consumo de conteúdo de jogos ao vivo por semana), “age” (idade do jogador), “work” (situação profissional), “degree” (nível educacional) e “playstyle” (estilo do jogo).

Essas variáveis foram escolhidas para identificar possíveis influências do estilo de vida do jogador, do tipo de jogo que está sendo consumido e do envolvimento da monetização do ato de jogar, permitindo que o modelo considere não apenas os sintomas autodeclarados, e sim, hábitos e características relacionadas ao tema que podem ter impacto no bem estar do jogador.

4.4 Construção da rede neural

4.4.1 Arquitetura do modelo

Com os dados já tratados, a modelagem foi conduzida por meio de três modelos de classificação em Redes Neurais Densas (DNN). A escolha desse tipo de rede deve-se à sua capacidade de aproximação universal e sua flexibilidade para formar soluções de qualidade para uma ampla classe de problemas, segundo (Zuben, 1993), utilizando a *API Sequential* da biblioteca *keras* (Chollet, 2025).

Todas as modelagens seguiram uma mesma estrutura metodológica, variando apenas

quanto à variável-alvo e ao número de classes envolvidas em cada tarefa.

Durante a definição da arquitetura das redes, optou-se por utilizar a função de ativação ReLU (*Rectified Linear Unit*) nas camadas ocultas por suas vantagens comprovadas em redes profundas. Estudos como (Dubey; Singh; Chaudhuri, 2021) demonstraram que a função ReLU tem uma alta taxa de uso em redes profundas.

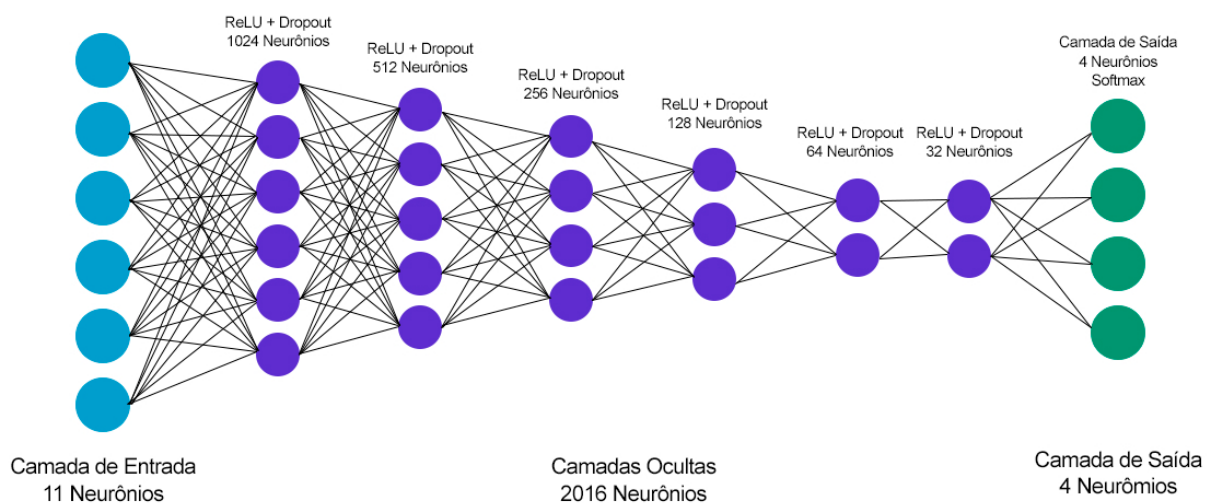
Para a camada de saída, foi adotado o uso da função de ativação Softmax, por sua capacidade de gerar uma distribuição de probabilidade normalizada sobre as classes previstas. Além disso, o Softmax combinado à *Cross-Entropy* é uma boa abordagem para classificação multiclases em redes neurais. O estudo (Li et al., 2020), demonstra que a *Softmax Loss* resolve problemas de classificação multiclases ao considerar todas as classes simultaneamente, promovendo mais estabilidade na camada de saída do modelo.

A arquitetura das redes foi composta por diversas camadas densas, contendo seis camadas de processamento, possuindo cerca de dois mil neurônios com taxa de *dropout* em 20% para mitigar o *overfitting*, e a camada de saída varia conforme a variável de interesse.

- **GAD-7:** quatro classes (Mínima, Leve, Moderada e Severa);
- **SPIN:** cinco classes (Nulo, Leve, Moderado, Severo, Muito Severo);
- **SWL:** sete classes (Muito Insatisfeito, Insatisfeito, Pouco Insatisfeito, Neutro, Pouco Satisfeito, Satisfeito, Muito Satisfeito.).

Com base nessa configuração, a Figura 14 ilustra a arquitetura final das redes implementadas para cada uma dos índices analisados.

Figura 14 – Arquitetura da Rede Neural Densa responsável pelo índice GAD-7



Embora a arquitetura das redes varie apenas no número de neurônios da camada de saída, todas compartilham exatamente a mesma estrutura de entrada e camadas ocultas. A diferença decorre exclusivamente da quantidade de classes de cada índice clínico: no GAD-7, a saída possui 4 neurônios, no SWL, 7 neurônios, e no SPIN, 5 neurônios.

4.4.2 Configuração do treinamento

O modelo foi compilado utilizando o otimizador Adam, por causa da sua eficiência e adaptabilidade (Kingma; Ba, 2014), e a função de perda utilizada foi a `SparseCategoricalCrossentropy` (TENSORFLOW, 2025), adequada para classificações com rótulos inteiros, evitando a necessidade de conversão para *one-hot encoding* (vetores binários), no fim, a métrica de desempenho do treinamento e validação, foi a acurácia dos resultados (*accuracy*).

Ademais, durante o treinamento, foi aplicado o *early stopping*, para que após 10 épocas com perda na validação, o treinamento seja interrompido para manter os melhores pesos do modelo. O treinamento teve o limite máximo de 150 épocas, com 20% deles sendo usados para validação.

4.5 Avaliação do modelo

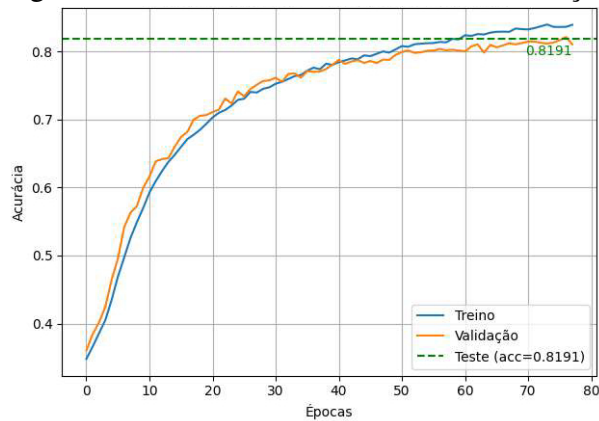
Os gráficos de treinamento apresentam a evolução das métricas de desempenho ao longo das épocas, tanto para o conjunto de treino quanto para o de validação. Essa visualização permite identificar fenômenos como convergência, estabilidade, sobreajuste ou subajuste, ajudando a avaliar se o modelo aprendeu padrões relevantes ou se memorizou os dados.

A performance preditiva do modelo foi avaliada com base na acurácia (proporção de classificações corretas). Os resultados foram obtidos a partir do conjunto de teste, o que permitiu estimar a generalização do modelo para novos dados.

Para cada tarefa de classificação (GAD-7, SWL e SPIN), foi produzido um gráfico de desempenho por época (Figura 15, Figura 16 e Figura 17, respectivamente), comparando a acurácia obtida nos dados de treino e validação, possibilitando uma análise visual da convergência e estabilidade do treinamento, bem como a distância entre os desempenhos nos conjuntos, servindo de apoio à identificação de possíveis indícios de *overfitting* e *underfitting*.

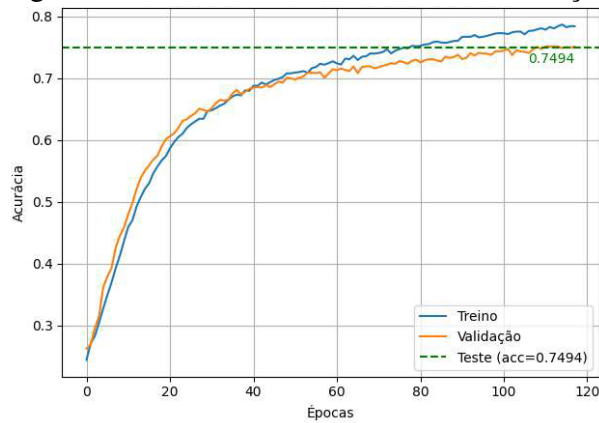
A acurácia final do teste foi destacada graficamente por uma linha horizontal verde, facilitando a visualização do ponto de referência para comparação de treino e teste.

Figura 15 – Gráfico de acurácia da classificação por época para o índice GAD-7



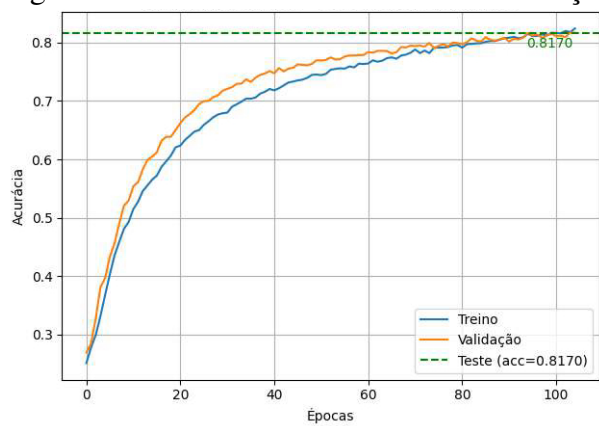
Fonte: Autoria Própria (2025).

Figura 16 – Gráfico de acurácia da classificação por época para o índice SWL



Fonte: Autoria Própria (2025).

Figura 17 – Gráfico de acurácia da classificação por época para o índice SPIN



Fonte: Autoria Própria (2025).

5 RESULTADOS

5.1 Acurácia final de cada rede

Os valores demonstram a capacidade dos modelos em aprender os padrões comportamentais associados a cada índice clínico. Observa-se que, apesar de pequenas oscilações, os modelos mantiveram estabilidade, convergindo com diferença mínima entre treino e validação, o que indica ausência de *overfitting* acentuado, especialmente observado no SWL, onde, à partir da 80ª época, o *early stopping* foi acionado para evitar perdas.

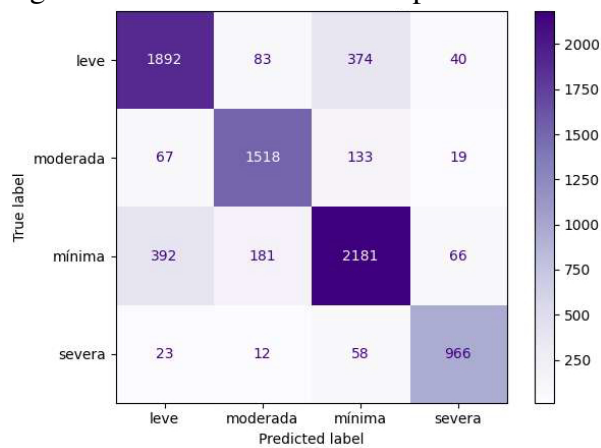
5.2 Matriz de confusão e análise dos erros

As matrizes de confusão permitem visualizar o desempenho do modelo ao comparar as classes previstas com as classes reais. Cada célula representa a quantidade de observações classificadas em uma combinação específica de rótulo verdadeiro e rótulo previsto, o que possibilita identificar acertos, erros e padrões de confusão entre classes. Esse tipo de visualização é especialmente útil para avaliar o comportamento do modelo em cenários de desequilíbrio entre classes.

As matrizes de confusão (Figura 18, Figura 19 e Figura 20) ilustram a performance dos modelos por classe e são essenciais para esta análise:

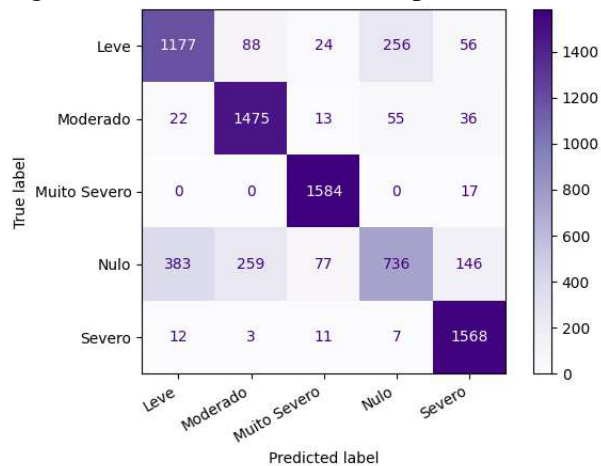
- O modelo GAD-7 (Figura 18) apresenta maior precisão nas classificações extremas (mínima e severa), com confusões pontuais entre os níveis leve e moderado.
- No SPIN (Figura 19), a classe nulo apresentou maior taxa de confusão, sendo frequentemente classificada como níveis leves ou moderados de fobia social. Isso indica que o modelo é sensível à presença de sintomas, mas pode superestimar sinais em indivíduos sem manifestações clínicas.
- Para o modelo SWL (Figura 20), os erros se concentram nos níveis adjacentes (por exemplo, entre “pouco satisfeito” e “neutro”), o que é compreensível, dada a subjetividade da escala.

Figura 18 – Matriz de confusão para o índice GAD-7



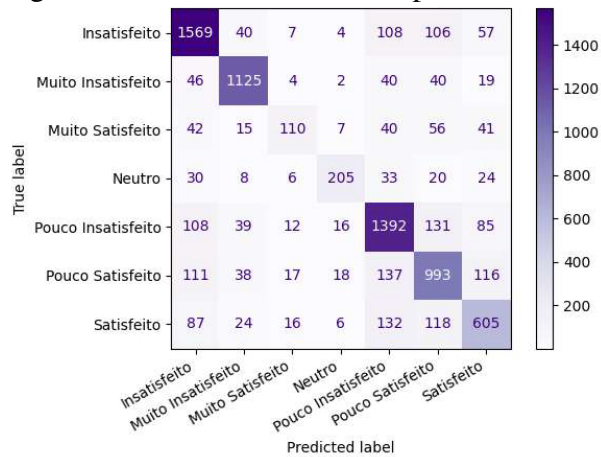
Fonte: Autoria Própria (2025).

Figura 19 – Matriz de confusão para o índice SPIN



Fonte: Autoria Própria (2025).

Figura 20 – Matriz de confusão para o índice SWL



Fonte: Autoria Própria (2025).

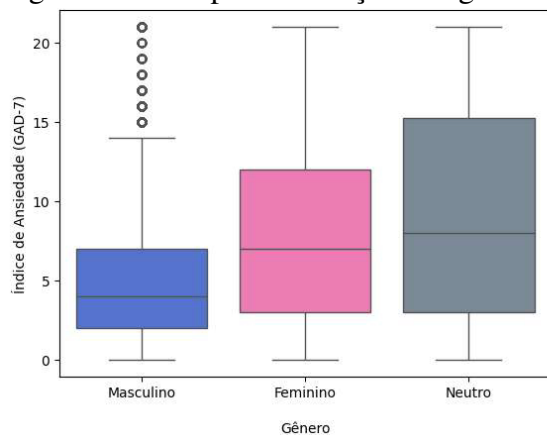
Apesar disso, a rede demonstrou habilidade em distinguir extremos (como "muito satisfeito" versus "muito insatisfeito").

5.3 Distribuição de características por índice clínico

Análises adicionais foram realizadas para compreender a distribuição das pontuações “GAD-7”, “SWL”, “SPIN” com base em variáveis como gênero e tempo de jogo.

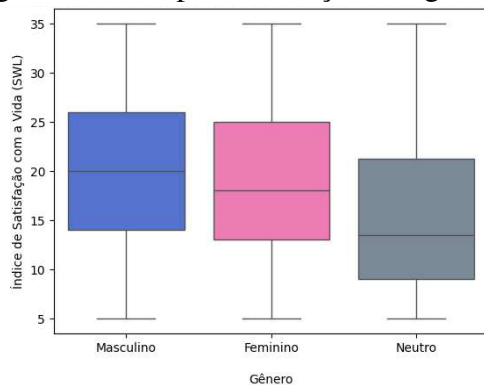
Os boxplots sintetizam a dispersão dos dados ao representar quartis, medianas e possíveis valores atípicos de cada grupo analisado. Esse tipo de gráfico facilita a comparação entre categorias, permitindo identificar diferenças de variabilidade, assimetria e concentração, além de revelar possíveis padrões relevantes entre variáveis comportamentais e níveis clínicos.

Figura 21 – Boxplot de relação dos gêneros com a média do índice GAD-7



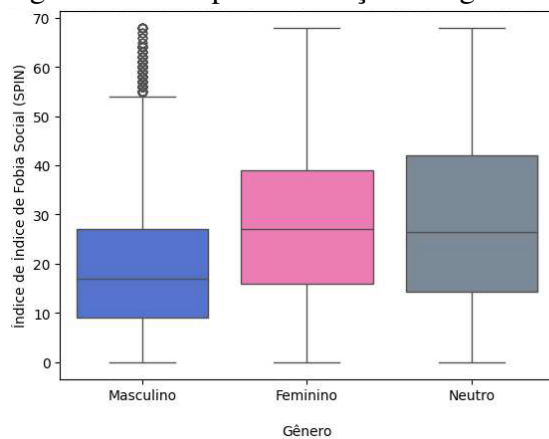
Fonte: Autoria Própria (2025).

Figura 22 – Boxplot de relação dos gêneros com a média do índice SWL



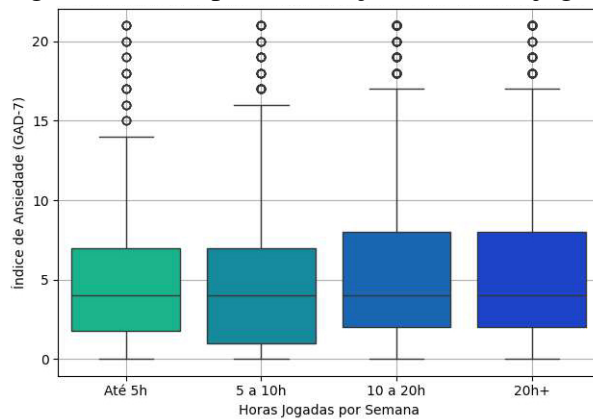
Fonte: Autoria Própria (2025).

Figura 23 – Boxplot de relação dos gêneros com a média do índice SPIN



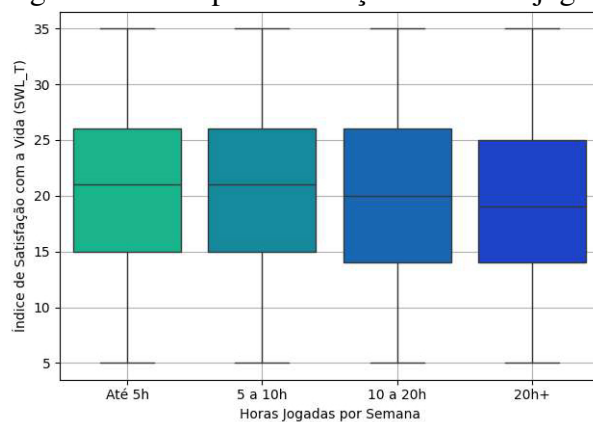
Fonte: Autoria Própria (2025).

Figura 24 – Boxplot de relação das horas jogadas com a média do índices clínico GAD-7



Fonte: Autoria Própria (2025).

Figura 25 – Boxplot de relação das horas jogadas com a média do índices clínico SWL



Fonte: Autoria Própria (2025).

Os *Boxplots* das Figuras 21, 22, 23 indicam que indivíduos do sexo feminino apresentaram medianas ligeiramente maiores nos escores de ansiedade e fobia social. Além disso, usuários que

jogam mais de 5 horas diárias apresentaram tendência a maiores níveis de ansiedade e menor satisfação com a vida (Figura 24 e Figura 25). Esses achados reforçam a hipótese de que os comportamentos nos jogos estão fortemente associados ao estado psicológico dos indivíduos.

5.4 Perfis típicos previstos

Com base nos dados de entrada que levaram a cada classificação, foi possível observar perfis típicos por faixa clínica. Esses perfis foram construídos a partir da média dos atributos comportamentais dos indivíduos que, segundo a rede neural, pertenciam a cada nível clínico.

As tabelas geradas (Figura 26, Figura 27 e Figura 28), sintetizam as principais características desses grupos, considerando as variáveis comportamentais exploradas anteriormente na modelagem dos perfis.

Figura 26 – Perfis previstos pela rede baseado no índice GAD-7

	Idade Média	Playstyle	Game	Platform	whyplay	Work	Degree	League	earnings	Horas por Semana	Streams por Semana
mínima	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
leve	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
moderada	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	10
severa	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9

Fonte: Autoria Própria (2025).

Os resultados obtidos pela rede neural acompanham tendências identificadas na literatura. Estudos como o de Vliet (2022) apontam que a maior parte dos jogadores tende a apresentar níveis mínimos ou leves de ansiedade, enquanto níveis moderados e severos são proporcionalmente menos frequentes, padrão semelhante ao observado neste trabalho. De modo compatível, Zhao et al. (2021) destacam que sintomas de ansiedade mais intensos são menos comuns entre jovens jogadores, embora associados a maior vulnerabilidade emocional e padrões de uso digital mais desgastantes. Assim, a distribuição prevista pelo modelo para o GAD-7 mantém coerência com achados prévios, reforçando que níveis elevados de ansiedade, ainda que presentes, constituem uma parcela menor da amostra, mas permanecem relevantes para análises psicológicas aplicadas ao comportamento em ambientes digitais.

Figura 27 – Perfis previstos pela rede baseado no índice SWL

	Idade Média	Playstyle	Game	Platform	whyplay	Work	Degree	League	earnings	Horas por Semana	Streams por Semana
Insatisfeito	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
Muito Insatisfeito	21	Outro	Diablo 3	Console	Diversão	Outro	Graduação	Diamond	Diversão + Renda parcial	21	10
Muito Satisfeito	21	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	10
Neutro	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	18	9
Pouco Insatisfeito	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
Pouco Satisfeito	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	19	9
Satisfeito	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	19	9

Fonte: Autoria Própria (2025).

A predominância de níveis insatisfeitos ou pouco satisfeitos prevista para o índice SWL é consistente com os achados de Vliet (2022), que identificaram associação entre menor satisfação com a vida, sintomas de ansiedade e padrões elevados de engajamento digital. Estudos mais recentes, como Zhao et al. (2021), também relatam que comportamentos de uso excessivo de jogos podem estar relacionados à redução de indicadores de bem-estar subjetivo, sobretudo entre populações jovens. Dessa forma, os resultados produzidos pelo modelo mantêm alinhamento com as evidências existentes ao indicar que o jogo pode funcionar como estratégia momentânea de regulação emocional, mas não garante, necessariamente, níveis mais altos de satisfação global com a vida.

Figura 28 – Perfis previstos pela rede baseado no índice SPIN

	Idade Média	Playstyle	Game	Platform	whyplay	Work	Degree	League	earnings	Horas por Semana	Streams por Semana
Leve	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
Moderado	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9
Muito Severo	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	10
Nulo	21	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	18	9
Severo	20	Outro	Diablo 3	Console	Diversão	Empregado	Graduação	Diamond	Diversão + Renda parcial	20	9

Fonte: Autoria Própria (2025).

A distribuição prevista para os níveis de fobia social (avaliados pelo SPIN) também dialoga com a literatura especializada. A revisão realizada por Gioia, Colella e Boursier (2022) demonstra que sintomas leves e moderados de ansiedade social são os mais recorrentes entre jogadores de ambientes competitivos, enquanto níveis severos aparecem com menor frequência. Esse comportamento é compatível com os dados observados no presente estudo. Resultados semelhantes são descritos por (Divyansh22, 2021), que identifica incidência maior de sintomas

leves de ansiedade social entre usuários de jogos online. Dessa forma, os padrões previstos pela rede neural reforçam que a fobia social tende a manifestar-se de maneira graduada, com maior concentração em níveis mais baixos ou intermediários.

Em geral, observou-se relativa homogeneidade em alguns atributos, como: idade média e plataforma preferida; enquanto outros, como estilo de jogo e liga competitiva, apresentaram maior variação entre os níveis clínicos.

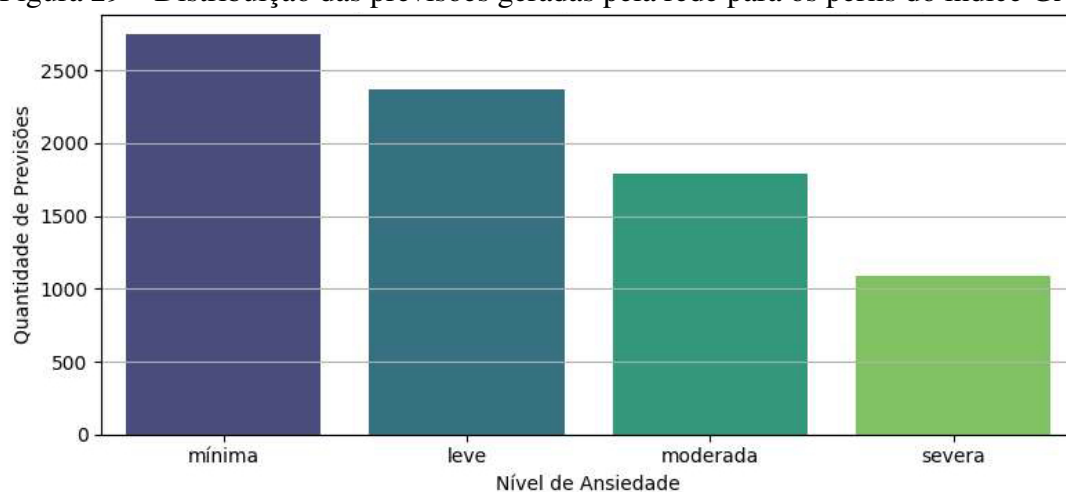
Cabe destacar que, a presença de categorias dominantes em determinados níveis, pode refletir tendências específicas do público da amostra, mas também pode indicar limitações dos dados.

5.5 Distribuição das previsões por nível clínico

Além da análise dos perfis médios previstos, também foi avaliada a quantidade das classificações atribuídas pela rede neural para cada faixa clínica dos índices GAD-7, SWL e SPIN. Essa análise, visa demonstrar eventuais desequilíbrios nas previsões e na compreensão da frequência com que certos padrões são associados a níveis de ansiedade, fobia social ou satisfação com a vida.

A Figura 29, apresenta a distribuição das previsões de ansiedade (GAD-7). Nota-se que, a maior parte dos indivíduos foi classificada nos níveis mínima e leve, com números progressivamente menores nos níveis moderada e severa.

Figura 29 – Distribuição das previsões geradas pela rede para os perfis do índice GAD-7

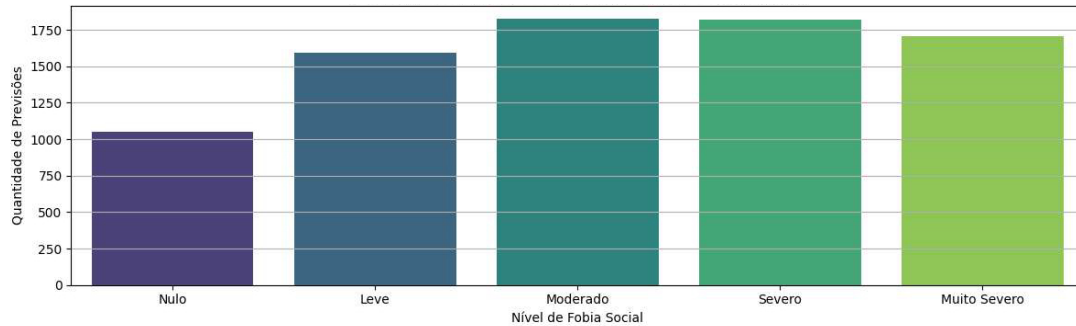


Fonte: Autoria Própria (2025).

A Figura 30, exibe as classificações previstas para os níveis de satisfação com a vida

(SWL), com destaque para a predominância dos níveis insatisfeitos e pouco satisfeitos. Os níveis com maior satisfação, foram menos frequentes nas previsões.

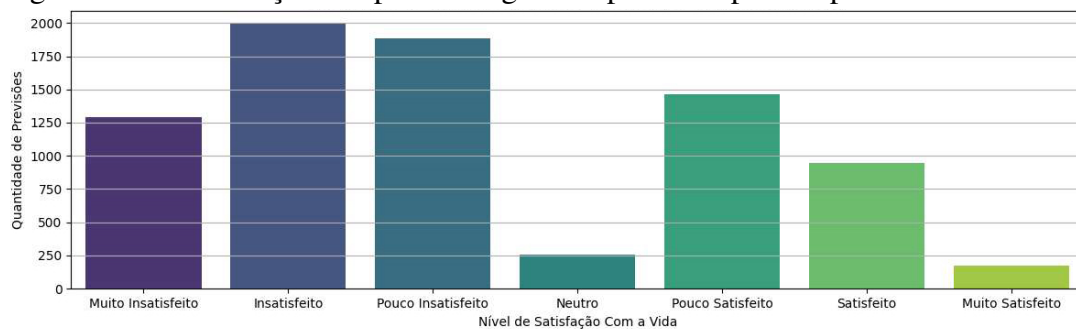
Figura 30 – Distribuição das previsões geradas pela rede para os perfis do índice SWL



Fonte: Autoria Própria (2025).

Por fim, a Figura 31, mostra a distribuição das previsões de fobia social (SPIN). Percebe-se um equilíbrio maior entre as categorias leve, moderado, severo e muito severo; enquanto a categoria nulo apresentou menor frequência.

Figura 31 – Distribuição das previsões geradas pela rede para os perfis do índice SPIN



Fonte: Autoria Própria (2025).

Esses dados reforçam a importância de considerar a distribuição das classes previstas ao interpretar os perfis típicos, especialmente em contextos com viés comportamental ou psicológico.

5.6 Ameaças à validade e limitações

A interpretação dos resultados deste estudo exige a consideração de algumas ameaças à validade, decorrentes tanto das características da base quanto das decisões metodológicas adotadas. A natureza auto-relatada dos dados representa uma ameaça à validade de construto, pois respostas subjetivas podem introduzir vieses de percepção, exagero ou subestimação de

sintomas. Ainda que instrumentos como GAD-7, SPIN e SWL sejam amplamente validados, sua aplicação fora de contextos clínicos controlados pode gerar variações não previstas nos instrumentos originais.

A imputação pela média nas variáveis numéricas, utilizada devido à baixa proporção de valores ausentes, preserva o formato geral das distribuições, mas pode reduzir parcialmente a variabilidade dos atributos, afetando a validade interna. Alternativas mais complexas de imputação poderiam gerar resultados ligeiramente diferentes, embora também pudessem introduzir padrões artificiais que comprometeriam o aprendizado.

O desequilíbrio entre classes, mesmo após o oversampling, constitui outra ameaça, pois modelos preditivos tendem a apresentar desempenho inferior em categorias naturalmente menos frequentes. Assim, as métricas para classes minoritárias podem superestimar a capacidade real de generalização. Esse risco é amplificado pela natureza específica da população estudada (jogadores online), o que limita a validade externa e impede a generalização direta dos achados para populações clínicas amplas ou grupos demográficos distintos.

Por fim, a escolha por utilizar redes neurais densas, embora adequada ao objetivo preditivo, implica assumir características de modelos de caixa-preta, o que restringe a interpretabilidade dos padrões aprendidos. Técnicas alternativas, como modelos lineares ou árvores de decisão, poderiam oferecer maior transparência, mas possivelmente com desempenho inferior na captura de relações não lineares.

Apesar dessas ameaças, os resultados obtidos permanecem válidos dentro do escopo definido, desde que interpretados considerando as condições experimentais e as limitações inerentes aos dados e às escolhas metodológicas adotadas.

6 CONCLUSÕES E TRABALHOS FUTUROS

6.1 Conclusões

Este estudo demonstrou o potencial de redes neurais multicamadas em prever níveis clínicos de ansiedade (GAD-7), fobia social (SPIN) e satisfação com a vida (SWL) com base em variáveis comportamentais autodeclaradas por jogadores. Os modelos apresentaram acurácia satisfatória, estabilidade durante o treinamento e capacidade interpretativa dos padrões de entrada.

Além de validar a metodologia proposta, os resultados indicam que há potencial para o uso dessas arquiteturas como suporte no rastreio precoce de condições psicológicas.

Do ponto de vista interpretativo, os padrões identificados pelas redes sugerem relações entre características comportamentais e os níveis clínicos previstos. Em contextos clínicos, tal ferramenta poderia acelerar o processo de triagem, que pode levar semanas até que um profissional consiga identificar com precisão o quadro do paciente, sobretudo em ambientes de alta demanda e poucos recursos.

Do ponto de vista social e acadêmico, o uso dessas ferramentas pode abrir caminho para novas pesquisas, por exemplo, o tempo médio que um psicólogo ou psiquiatra leva para diagnosticar formalmente um paciente com o tempo de resposta obtido via predição automatizada.

Importante destacar que, a proposta não tem o objetivo de substituir psicólogos, psiquiatras ou quaisquer profissionais da área, mas sim, de atuar como uma ferramenta de apoio, otimizando o processo de triagem e possibilitando uma alocação mais eficiente dos recursos humanos e do tempo clínico.

6.2 Trabalhos futuros

Como continuidade deste trabalho, propõe-se:

- **Inclusão de novos índices clínicos**

Para auxiliar na completude do modelo, ampliando a sua análise para outras dimensões dentro da saúde mental, a utilização de novos índices é uma sugestão a ser considerada, tendo como exemplos os seguintes índices:

- *Patient Health Questionnaire-9* (PHQ-9) - Gravidade da Depressão: Instrumento multifuncional que permite a triagem, diagnóstico, monitoramento e medição da gravidade da depressão de uma forma padronizada.

- *Internet Gaming Disorder Scale-Short Form* (IGDS9-SF) - Desordem de Jogos na Internet: Ele avalia sintomas que podem ser notados após a exposição problemática de jogos.
- *Pittsburgh Sleep Quality Index* (PSQI) - Qualidade do Sono: Teste que busca investigar a qualidade do sono, pois uma qualidade baixa pode ser um dos sintomas da ansiedade ou depressão.
- *Hamilton Depression Rating Scale* (HAM-D) - Severidade da Depressão: Desenvolvida por Max Hamilton, a escala permite avaliar sintomas de depressão em adultos, é frequentemente utilizada para avaliar a evolução do tratamento antidepressivo.
- *Hamilton Anxiety Rating Scale* (HAM-A) - Severidade da Ansiedade: Também criada por Max Hamilton, é uma escala que foca em medir a gravidade dos principais sintomas da ansiedade, utilizada no diagnóstico e acompanhamento para medir a resposta ao tratamento.

- **Testes com arquiteturas alternativas**

A proposta é de utilizar arquiteturas novas como as redes convolucionais 1D, que podem captar padrões locais e sequenciais dos dados, ou redes *transformers* adaptadas para possibilitar a modelagem de dependências ao longo do prazo entre as variáveis analisadas, ampliando o potencial preditivo.

- **Exploração de técnicas de *Explainable AI***

Métodos como *Local Interpretable Model-Agnostic Explanations* (LIME) ou *Shapley Additive Explanations* (SHAP), podem ser aplicados para interpretar o impacto de cada variável de entrada nas decisões do modelo, oferecendo explicações mais transparentes e compreensíveis para o auxílio no diagnóstico.

REFERÊNCIAS

- ALI, H. et al. A review on data preprocessing methods for class imbalance problem. **International Journal of Engineering & Technology**, v. 8, n. 3, p. 390–397, 2019. Disponível em: <https://www.researchgate.net/publication/336305785>.
- AMORIM, T. et al. The choice of scaling technique matters for classification performance. **Applied Soft Computing**, 2022. Disponível em: <https://doi.org/10.1016/j.asoc.2022.109924>.
- ARRIETA, A. B. et al. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. **Information Fusion**, v. 58, p. 82–115, 2020. ISSN 1566-2535. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1566253519308103>.
- AZEVEDO, J. C.; NASCIMENTO, G. do; SOUZA, C. H. M. de. Ciberdependência: o papel das emoções na dependência de tecnologias digitais. **Texto Livre: Linguagem e Tecnologia**, v. 7, n. 2, p. 124–141, 2014. Disponível em: RedALyC. Acesso em: 25 jul. 2025. Disponível em: <https://www.redalyc.org/articulo.oa?id=577163625002>.
- BARRETO, L. S. d. B. **Apostas on line: o outro lado do jogo**. Goiânia: Pontifícia Universidade Católica de Goiás, 2025. 39 f. Monografia Bacharelado em Jornalismo). Disponível em: <https://repositorio.pucgoias.edu.br/jspui/handle/123456789/9311>.
- BARTLE, R. Hearts, clubs, diamonds, spades: Players who suit muds. **Journal of MUD Research**, v. 1, n. 1, p. 1–26, 1996. Disponível em: <https://mud.co.uk/richard/hcds.htm>.
- BRAGA, A. d. P.; LUDERMIR, T. B.; CARVALHO, A. C. P. d. L. F. d. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007. ISBN 978-8521615644.
- CHOLLET, F. **Keras: Biblioteca de aprendizado profundo em Python**. 2025. <https://keras.io/>. Acesso em: 25 jul. 2025.
- CONNOR, K. M. et al. Psychometric properties of the social phobia inventory (spin): new self-rating scale. **British Journal of Psychiatry**, v. 176, n. 4, p. 379–386, 2000. Disponível em: <https://callhelpline.org.uk/app/uploads/2024/03/Social-Phobia-Inventory-SPIN-Form.pdf>.
- DIENER, E. et al. **Satisfaction with life scale (SWLS): instructions and scale**. 2024. NovoPsych. Disponível em: https://novopsych.com/wp-content/uploads/2024/11/swls_satisfaction-with-life-scale_assessment.pdf. Updated PDF version.
- DIVYANSH22. **Online Gaming Anxiety Data**. 2021. <https://www.kaggle.com/datasets/divyansh22/online-gaming-anxiety-data>. Kaggle. Acesso em: 25 jul. 2025.
- DUBEY, S. R.; SINGH, S. K.; CHAUDHURI, B. B. Activation functions in deep learning: A comprehensive survey and benchmark. **arXiv preprint arXiv:2109.14545**, 2021. Disponível em: <https://arxiv.org/abs/2109.14545>.
- DUMAS, A. **O Conde de Monte Cristo**. Edição de bolso. Rio de Janeiro: Zahar, 2008. Tradução brasileira.
- GIOIA, F.; COLELLA, G. M.; BOURSIER, V. Evidence on problematic online gaming and social anxiety over the past ten years: a systematic literature review. **Current Addiction Reports**, v. 9, 2022. Disponível em: <https://doi.org/10.1007/s40429-021-00406-3>.

GUO, W. et al. Deep neural networks for depression recognition based on 2d and 3d facial expressions under emotional stimulus tasks. **Frontiers in Neuroscience**, Volume 15 - 2021, p. 609760, 2021. ISSN 1662-453X. Disponível em: <https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2021.609760>.

HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. Tradução da 2ª edição norte-americana. ISBN 978-8573077186.

HONG, S. J. et al. Severity identification for internet gaming disorder using heart rate variability reactivity for gaming cues: a deep learning approach. **Frontiers in Psychiatry**, Volume 14, 2023. ISSN 1664-0640. Disponível em: <https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2023.1231045>.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014. Disponível em: <https://arxiv.org/pdf/1412.6980>.

LEUNG, L.; LEE, P. S. N. Multiple determinants of life quality: The roles of internet activities, use of new media, social support, and leisure activities. **Telematics and Informatics**, v. 22, n. 3, p. 161–180, 2005. Disponível em: <https://doi.org/10.1016/j.tele.2004.04.003>.

LI, X. et al. On the learning property of logistic and softmax losses for deep neural networks. **arXiv preprint arXiv:2003.02309**, 2020. Disponível em: <https://arxiv.org/abs/2003.02309>.

MOLNAR, C. **Interpretable Machine Learning: A Guide for Making Black Box Models Explainable**. 2. ed. [S.l.]: Leanpub, 2022. Disponível em: <https://christophm.github.io/interpretable-ml-book/>.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (CID-11)**. 2019. <https://icd.who.int/en>. Acesso em: 10 dez. 2025.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Saúde mental: fortalecendo nossa resposta**. 2022. <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response>. Acesso em: 10 dez. 2025.

SAUTER, M.; DRASCHKOW, D. **Gaming Habits Dataset**. 2022. Open Science Framework (OSF). Disponível em: <https://osf.io/vnbxk/>.

SPITZER, R. L. et al. **GAD-7 Anxiety: Generalized Anxiety Disorder 7-item screener**. 2025. Anxiety and Depression Association of America. Disponível em: https://adaa.org/sites/default/files/GAD-7_Anxiety-updated_0.pdf.

TABARES, M. T. et al. Anxiety in young people: Analysis from a machine learning model. **Acta Psychologica**, v. 248, p. 104410, 2024. ISSN 0001-6918. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0001691824002877>.

TENSORFLOW. **SparseCategoricalCrossentropy — TensorFlow Core v2 API**. 2025. https://www.tensorflow.org/api_docs/python/tf/keras/losses/SparseCategoricalCrossentropy. Acesso em: 25 jul. 2025.

VLIET, T. van. **Predicting the Life Satisfaction of Online Gamers with Neural Networks**. Dissertação (Mestrado) — Tilburg University, School of Humanities and Digital Sciences, Department of Cognitive Science and Artificial Intelligence, Tilburg, The Netherlands, 2022. Disponível em: <https://arno.uvt.nl/show.cgi?fid=161835>.

WANG, Q. et al. Deep neural network to differentiate internet gaming disorder from healthy controls during stop-signal task: a multichannel near-infrared spectroscopy study. **Biomedical Engineering / Biomedizinische Technik**, v. 68, n. 5, p. 457–468, 2023. Disponível em: <https://doi.org/10.1515/bmt-2023-0030>.

YEE, N. Motivations for play in online games. **CyberPsychology & Behavior**, v. 9, n. 6, p. 772–775, 2006. Disponível em: <https://doi.org/10.1089/cpb.2006.9.772>.

YOUNAS, R.; HAQ, H. Burhan ul. Deep learning model to analyze psychological effects on gamers. **Spectrum of Decision Making and Applications**, v. 2, n. 1, p. 120–134, Jan. 2025. Disponível em: <https://dmap-journal.org/index.php/dmap/article/view/4>.

ZHAO, W. et al. The influence of online game behaviors on the emotional state and executive function of college students in china. **Frontiers in Psychiatry**, v. 12, p. 713364, 2021. ISSN 1664-0640. Disponível em: <https://doi.org/10.3389/fpsy.2021.713364>.

ZUBEN, F. J. V. **Redes neurais artificiais**. Dissertação (Mestrado) — Universidade Estadual de Campinas (Unicamp), Departamento de Engenharia de Computação e Automação, Campinas, SP, Brasil, 1993. Cap. 2 da tese de mestrado. Disponível em: https://www.dca.fee.unicamp.br/von-zuben/theses/emi_mest/cap2.pdf.

ŞİPOŞ, R. et al. Analysing digital engagement patterns: A machine learning investigation into social anxiety among adolescents with adhd. **Journal of Clinical Medicine**, v. 13, n. 23, 2024. ISSN 2077-0383. Disponível em: <https://www.mdpi.com/2077-0383/13/23/7461>.