



**Universidade Federal do Maranhão**  
**Centro de Ciências Exatas e Tecnologia**  
**Programa de Pós-Graduação em Engenharia Elétrica**

**Sistema de Apoio ao Diagnóstico do Risco de Transtorno de  
Ansiedade em Crianças**

**Renata Costa Rocha**

**São Luís – MA**  
**2026**

**Universidade Federal do Maranhão**  
**Centro de Ciências Exatas e Tecnologia**  
**Programa de Pós-Graduação em Engenharia Elétrica**

**Renata Costa Rocha**

**Sistema de Apoio ao Diagnóstico do Risco De Transtorno de  
Ansiedade em Crianças**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, do Centro de Ciências Exatas e Tecnologia da Universidade Federal do Maranhão, em cumprimento às exigências institucionais para obtenção do grau de Mestre em Engenharia Elétrica, área de Automação e Controle, sob orientação do Prof. Dr. Allan Kardec Dualibe Barros Filho, e Coorientação do Prof. Dr. Jonathan Araújo Queiroz.

**São Luís – MA**

**2026**

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).  
Diretoria Integrada de Bibliotecas/UFMA

Costa Rocha, Renata.

Sistema de Apoio ao Diagnóstico do Risco de Transtorno de Ansiedade em Crianças / Renata Costa Rocha. - 2026.  
85 f.

Coorientador(a) 1: Jonathan Araújo Queiroz.

Orientador(a): Allan Kardec Dualibe Barros Filho.

Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia Elétrica/ccet, Universidade Federal do Maranhão, São Luís, 2026.

1. Vetores de Atributos Psicométricos. 2. Ansiedade Pediátrica. 3. Classificação Multinível. 4. Apoio À Decisão Clínica. I. Dualibe Barros Filho, Allan Kardec. II. Araújo Queiroz, Jonathan. III. Título.

*“Corrida hoje, vitória amanhã.”*

***Racionais MC's***

*Dedico esta Dissertação de Mestrado primeiramente a Deus, onipresente em todas as minhas batalhas; aos meus filhos Henrique e Yasmin, que são o melhor de mim e razão maior da minha força e dedicação; ao meu esposo Yuri, companheiro e amor da vida.*

*Agradeço a Deus, pelo discernimento e pela força concedidos; ao meu esposo Yuri e aos meus filhos Henrique e Yasmin, que com paciência e amor me ofereceram suporte e foram meu porto seguro; às crianças que participaram do estudo de Harvard, cuja valiosa contribuição possibilitou a criação do banco de dados aberto e, consequentemente, o desenvolvimento deste trabalho; ao meu orientador, Prof. Dr. Allan Kardec Dualibe Barros Filho, pelas orientações, ensinamentos e direcionamentos ao longo desta jornada; e ao meu coorientador, Prof. Dr. Jonathan Queiroz, pela colaboração, apoio acadêmico e contribuições fundamentais para a condução e consolidação deste estudo.*

## RESUMO

Os transtornos de ansiedade na infância configuram um problema relevante de saúde mental, particularmente em crianças em idade pré-escolar, devido à dificuldade de identificação precoce associada à subjetividade dos instrumentos clínicos e à sobreposição de padrões comportamentais. Nesse contexto, abordagens computacionais têm sido investigadas como ferramentas de apoio à decisão clínica, especialmente aquelas baseadas em técnicas de aprendizado de máquina aplicadas a dados psicométricos. Entretanto, tais abordagens ainda enfrentam desafios relacionados à robustez decisória, à interpretabilidade clínica e à consistência do desempenho em cenários multiclasse ordinais. Com o objetivo de enfrentar essas limitações, este trabalho propõe um sistema computacional de apoio à decisão para a classificação binária e multinível do risco de transtornos de ansiedade em crianças, fundamentado em modelos de aprendizado supervisionado sob a perspectiva da Engenharia Elétrica aplicada à saúde mental. São avaliados comparativamente três modelos consolidados — *Random Forest*, *Support Vector Machine* e *Multilayer Perceptron* — considerando um cenário binário auxiliar de triagem clínica e um cenário multiclasse ordinal composto por quatro níveis de risco. A metodologia emprega uma base de dados pública do *Harvard Dataverse*, composta por informações psicométricas, comportamentais e psicossociais, sendo validada por meio de validação cruzada estratificada e métricas de desempenho alinhadas à prática clínica, incluindo análise explícita do erro clínico ordinal por meio da distinção entre erros adjacentes e não adjacentes no cenário multiclasse. Os resultados indicam que os modelos apresentam elevada capacidade discriminativa no cenário binário (AUC superiores a 0,94) e desempenho consistente no cenário multiclasse (acurácia global entre 83% e 85%), sem a existência de um classificador universalmente superior. Observa-se comportamento complementar entre as arquiteturas, com destaque para o *RF* em termos de estabilidade global e para o *MLP* na distinção entre níveis intermediários de risco. Conclui-se que modelos supervisionados clássicos, quando validados, constituem ferramentas promissoras para a estratificação ordinal do risco de ansiedade pediátrica. Sob a ótica da Engenharia Elétrica, o problema é caracterizado como um processo de inferência de estados latentes a partir de dados ruidosos e parcialmente observáveis, contribuindo metodologicamente para o desenvolvimento de sistemas computacionais interpretáveis de apoio à decisão em saúde mental infantil.

**Palavras-chave:** Vetores de Atributos Psicométricos. Ansiedade Pediátrica. Classificação Multinível. Apoio à Decisão Clínica.

# ABSTRACT

Childhood anxiety disorders represent a significant mental health concern, particularly due to the challenges associated with early identification, the subjectivity of clinical assessment instruments, and the overlap of behavioral patterns. In this context, computational approaches have been investigated as decision-support tools, especially those based on machine learning techniques applied to psychometric data. However, such approaches still face challenges related to decision robustness, clinical interpretability, and consistent performance in ordinal multiclass scenarios. To address these limitations, this work proposes a computational decision-support system for binary and multilevel classification of anxiety disorder risk in children, grounded in supervised learning models from the perspective of Electrical Engineering applied to mental health. Three established models are comparatively evaluated — *Random Forest*, *Support Vector Machine*, and *Multilayer Perceptron* — considering a binary screening scenario and an ordinal multiclass scenario composed of four risk levels. The methodology employs a public dataset from the *Harvard Dataverse*, consisting of psychometric, behavioral, and psychosocial information, and is validated through stratified cross-validation and performance metrics aligned with clinical practice, including explicit analysis of ordinal clinical error by distinguishing adjacent and non-adjacent misclassifications. The results indicate that the models achieve high discriminative capacity in the binary scenario (AUC values above 0.94) and consistent performance in the multiclass scenario (overall accuracy between 83% and 85%), with no universally superior classifier. A complementary behavior among the architectures is observed, with *Random Forest* demonstrating greater global stability and *Multilayer Perceptron* showing improved discrimination between intermediate risk levels. It is concluded that classical supervised models, when properly validated, constitute promising tools for the ordinal stratification of pediatric anxiety risk. From an Electrical Engineering perspective, the problem is characterized as a latent state inference process based on noisy and partially observable data, contributing methodologically to the development of interpretable computational decision-support systems in child mental health.

**Keywords:** Psychometric Feature Vectors. Pediatric Anxiety. Multilevel Classification. Clinical Decision Support.

## LISTA DE FIGURAS

Figura 1– Representação conceitual da arquitetura de um <i>Multilayer Perceptron (MLP)</i>	20
Figura 2– Representação conceitual do princípio de separação por <i>Support Vector Machines (SVM)</i>	22
Figura 3– Representação conceitual da estrutura de um <i>Random Forest (RF)</i>	24
Figura 4– Quadro comparativo das arquiteturas utilizadas no estudo	25
Figura 5– Representação conceitual da Matriz de Confusão	29
Figura 6– Relação conceitual entre Precisão, Sensibilidade e <i>F1-Score</i>	32
Figura 7– Curva <i>ROC</i> e área sob a curva ( <i>AUC</i> )	33
Figura 8– Representação conceitual validação cruzada estratificada	35
Figura 9– Estrutura geral da metodologia proposta	39
Figura 10– Matriz de Confusão Binária do modelo <i>SVM</i> (0 vs. 1–3)	49
Figura 11– Matriz de Confusão Binária do modelo <i>MLP</i> (0 vs. 1–3)	49
Figura 12– Matriz de Confusão Binária do modelo <i>RF</i> (0 vs. 1–3)	50
Figura 13– Curvas <i>ROC</i> – Classificação Binária (0 vs. 1–3)	51
Figura 14– Matriz de Confusão Multiclasse do modelo <i>SVM</i> (0–3)	53
Figura 15– Matriz de Confusão Multiclasse do modelo <i>MLP</i> (0–3)	53
Figura 16– Matriz de Confusão Multiclasse do modelo <i>RF</i> (0–3)	54
Figura 17– Importância Global das Variáveis no Modelo <i>MLP</i> ( <i>SHAP</i> Multiclasse)	55
Figura 18– Distribuição Percentual do Erro Clínico (Multiclasse)	57
Figura 19– Comparação de desempenho no cenário binário (0 vs 1–3)	59
Figura 20– <i>F1-score</i> por classe no cenário multiclasse (0–3)	60
Figura 21– Resumo comparativo multiclasse por gráfico radar	60
Figura 22– Resultados comparativos com estudos anteriores	67
Figura 23– Certificado de aceitação e apresentação oral no CILAMCE 2025	73
Figura 24– Certificado de aceitação e apresentação oral no CILAMCE 2025	73
Figura 25– Certificado de Registro de Programa de Computador expedido pelo INPI	74

## LISTA DE TABELAS

1	<b>Distribuição dos acertos e erros no cenário binário (0 vs 1–3).</b> . . . . .	48
2	<b><i>AUC (ROC) e AP (Precision–Recall) no cenário binário (0 vs 1–3).</i></b> . . . . .	51
3	<b>Acertos por classe no cenário multiclasse (0–3).</b> . . . . .	52
4	<b>Distribuição percentual do erro clínico no cenário multiclasse (0–3).</b> . . . . .	56
5	<b>Síntese das métricas de desempenho no cenário binário (0 vs 1–3).</b> . . . . .	58
6	<b><i>F1-score</i> por classe no cenário multiclasse (0–3).</b> . . . . .	59

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>5</b>
1.1	Estado da Arte . . . . .	6
1.2	Motivação . . . . .	8
1.3	Objetivos . . . . .	9
1.4	Organização do Trabalho . . . . .	10
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>11</b>
2.1	Transtorno de Ansiedade Pediátrica . . . . .	11
2.2	Psicometria como Sinal Discreto Ruidoso . . . . .	16
2.3	Redes Neurais Artificiais . . . . .	18
2.4	Métricas de Avaliação de Desempenho . . . . .	27
<b>3</b>	<b>METODOLOGIA</b>	<b>37</b>
3.1	Estrutura Geral da Metodologia . . . . .	38
3.2	Materiais . . . . .	39
3.3	Métodos . . . . .	44
<b>4</b>	<b>RESULTADOS</b>	<b>48</b>
4.1	Desempenho Binário (0 vs 1–3) . . . . .	48
4.2	Desempenho Multiclasse (0–3) . . . . .	52
4.3	Síntese dos Resultados . . . . .	58
4.4	Discussão dos Resultados . . . . .	62
<b>5</b>	<b>CONCLUSÃO E PERSPECTIVAS FUTURAS</b>	<b>69</b>
5.1	Limitações do Estudo . . . . .	70
5.2	Perspectivas Futuras . . . . .	71
	<b>Anexo A</b>	<b>72</b>
	<b>Anexo B</b>	<b>73</b>
	<b>Anexo C</b>	<b>74</b>
	<b>Referências</b>	<b>75</b>

# 1 INTRODUÇÃO

Os transtornos psiquiátricos com início na infância e na adolescência configuram-se como um problema de saúde pública de alcance global, apresentando elevada prevalência e impacto funcional significativo ao longo do desenvolvimento, conforme discutido por Vigo et al. [1]. Estimativas epidemiológicas indicam que uma parcela expressiva das crianças apresenta sintomas compatíveis com transtornos de ansiedade, os quais estão associados a prejuízos persistentes no desenvolvimento emocional, social e cognitivo, além de maior risco de psicopatologias na vida adulta, conforme relatado por Racine et al. [2] e Hense et al. [3].

Apesar da relevância clínica dessas condições, o paradigma diagnóstico predominante permanece fortemente fundamentado na semiologia psiquiátrica clássica, baseada em entrevistas clínicas, observação comportamental e relatos de pais e professores. Embora indispensáveis à prática clínica, tais instrumentos apresentam limitações inerentes, incluindo elevada subjetividade, variabilidade interobservador e dificuldade na identificação precoce e estratificação da gravidade dos sintomas, conforme apontado por Kuposov et al. [4]. Essas limitações tornam-se particularmente críticas em crianças em idade pré-escolar, cujas manifestações comportamentais frequentemente se sobrepõem a padrões típicos do desenvolvimento.

Nesse contexto, consolida-se o campo da psiquiatria computacional, uma área interdisciplinar que integra métodos quantitativos oriundos da Engenharia, da Neurociência e da Ciência de Dados com o objetivo de desenvolver ferramentas computacionais de apoio ao diagnóstico clínico baseadas em dados mensuráveis, conforme apresentado por Garcia et al. [5] e Shatte et al. [6]. Essa abordagem busca complementar a avaliação clínica tradicional por meio da modelagem matemática de processos subjacentes ao comportamento, à cognição e à regulação emocional.

Do ponto de vista da Engenharia Elétrica, o problema pode ser formulado como uma tarefa de processamento estatístico de sinais e reconhecimento de padrões, na qual variáveis psicométricas, comportamentais e psicofisiológicas constituem observações ruidosas de estados emocionais latentes. A extração de informação clinicamente relevante a partir desses dados demanda técnicas de pré-processamento, seleção de atributos e aprendizado supervisionado, capazes de capturar relações não lineares e padrões de alta complexidade, conforme descrito por Bishop [7] e por Hastie et al. [8].

Inserido nessa interface entre Engenharia Elétrica e saúde mental infantil, o presente trabalho propõe o desenvolvimento de um sistema computacional destinado à classificação binária

e multinível do risco de transtornos de ansiedade em crianças em idade pré-escolar, a partir de dados psicométricos e comportamentais estruturados em formato tabular, disponibilizados por Carpenter et al. [9] no *Harvard Dataverse*. A rotulagem diagnóstica da base fundamenta-se nos referenciais internacionais estabelecidos pelo Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5), publicado pela *American Psychiatric Association* [10], e pela Classificação Internacional de Doenças (CID-11), publicada pela *World Health Organization* [11].

Ao adotar modelos de aprendizado de máquina supervisionado amplamente consolidados — *Random Forest*, *Support Vector Machine* e *Multilayer Perceptron* — este estudo investiga a capacidade dessas arquiteturas em representar e discriminar níveis graduais de risco clínico, superando abordagens estritamente dicotômicas e alinhando-se à natureza dimensional dos transtornos ansiosos.

## 1.1 Estado da Arte

O uso de técnicas computacionais para apoio ao diagnóstico e à classificação de transtornos psiquiátricos tem se intensificado nas últimas décadas, impulsionado pelo avanço da capacidade computacional, pela crescente disponibilidade de bases de dados clínicas e pelo amadurecimento dos métodos de aprendizado de máquina. Sob a perspectiva da Engenharia Elétrica, esse campo insere-se no domínio do processamento estatístico de sinais e da modelagem de sistemas complexos, conforme discutido por Bishop [7], Hastie et al. [8] e Shatte et al. [6].

Estudos iniciais voltados à aplicação de aprendizado supervisionado em saúde mental concentraram-se predominantemente em esquemas de classificação binária, buscando distinguir indivíduos com e sem determinado transtorno. Algoritmos como *Support Vector Machines (SVM)*, propostos por Cortes e Vapnik [12], destacaram-se nesse contexto devido à sua capacidade de maximização de margem e desempenho robusto em espaços de alta dimensionalidade. Posteriormente, classificadores baseados em *Random Forest*, introduzidos por Breiman [13], passaram a ser amplamente adotados em aplicações clínicas, em virtude de sua robustez frente a ruído, dados desbalanceados e elevada variabilidade interindividual, conforme relatado por Shatte et al. [6] e Xu et al. [14].

No âmbito das redes neurais artificiais, arquiteturas do tipo *Multilayer Perceptron (MLP)* consolidaram-se como aproximadores universais de funções não lineares, conforme demonstrado por Hornik et al. [15] e discutido por Haykin [16]. Estudos recentes demonstram que redes *MLP* rasas, quando corretamente regularizadas, apresentam desempenho competitivo e

estabilidade superior em bases clínicas de tamanho moderado, em comparação com arquiteturas profundas mais complexas, conforme Lopes et al. [17] e Talaei-Khoei et al. [18].

Apesar desses avanços metodológicos, observa-se que grande parte da literatura permanece restrita a abordagens dicotômicas, nas quais o risco clínico é reduzido à presença ou ausência do transtorno. Tal simplificação contrasta com a natureza dimensional e contínua dos transtornos psiquiátricos, particularmente dos transtornos de ansiedade, cujas manifestações variam em intensidade, frequência e impacto funcional, conforme discutido por Stade et al. [19] e Caspi et al. [20]. A ausência de classificações multiníveis limita a aplicabilidade clínica dos modelos e dificulta a implementação de estratégias de triagem e intervenção proporcionais à gravidade do quadro.

No contexto específico da ansiedade pediátrica, essas lacunas tornam-se ainda mais evidentes. Estudos epidemiológicos recentes indicam elevada heterogeneidade fenotípica dos sintomas ansiosos e significativa sobreposição com outros transtornos do neurodesenvolvimento, como o Transtorno de Déficit de Atenção com Hiperatividade (TDAH), conforme relatado por Racine et al. [2] e Hense et al. [3]. Essa complexidade reforça a necessidade de modelos computacionais capazes de capturar relações não lineares e padrões estatísticos sutis, indo além de classificações binárias simplificadas.

Uma contribuição relevante nesse cenário foi apresentada por Carpenter et al. [9], que disponibilizaram uma base de dados pública no *Harvard Dataverse* contendo medidas psicométricas e comportamentais de crianças em idade pré-escolar, rotuladas de acordo com referenciais clínicos internacionais. Nesse estudo, os autores empregaram modelos simbólicos interpretáveis, como o *Alternating Decision Tree (ADTree)*, demonstrando o potencial de abordagens transparentes para a análise do risco de ansiedade pediátrica. Contudo, trabalhos posteriores destacam que tais modelos apresentam limitações de sensibilidade e generalização em contextos clínicos mais heterogêneos, conforme discutido por Xu et al. [14].

Pesquisas recentes têm explorado arquiteturas de maior complexidade, incluindo redes neurais profundas e modelos híbridos, sobretudo na análise de sinais fisiológicos contínuos, como eletroencefalografia (EEG) e variabilidade da frequência cardíaca (VFC), conforme Arif et al. [21] e Rivera et al. [22]. Entretanto, revisões sistemáticas recentes indicam que a aplicação dessas arquiteturas a dados psicométricos tabulares ainda carece de padronização metodológica e validação clínica consistente, conforme apontado por Talaei-Khoei et al. [18].

Outro aspecto de crescente relevância refere-se à interpretabilidade e explicabilidade

dos modelos. A incorporação de sistemas baseados em Inteligência Artificial em saúde mental impõe exigências éticas e regulatórias, demandando mecanismos que permitam compreender e justificar decisões automatizadas, conforme discutido por Shatte et al. [6] e Talia et al. [23]. Nesse contexto, modelos supervisionados clássicos e redes neurais rasas apresentam vantagens ao possibilitar análises de importância de atributos e integração com métodos de *Explainable Artificial Intelligence (XAI)*.

Em síntese, embora a literatura apresente avanços significativos na aplicação de aprendizado de máquina à saúde mental infantil, persistem lacunas relevantes relacionadas à predominância de classificações binárias, à limitada exploração de abordagens multiníveis em dados psicométricos tabulares e à escassez de estudos comparativos sistemáticos entre diferentes arquiteturas supervisionadas. O presente trabalho insere-se nesse contexto ao investigar, de forma comparativa e metodologicamente rigorosa, modelos supervisionados consolidados — *Random Forest*, *Support Vector Machine* e *Multilayer Perceptron* — aplicados à classificação multinível do risco de ansiedade pediátrica, utilizando dados públicos e alinhando-se às demandas contemporâneas da Engenharia Elétrica e da psiquiatria computacional.

## 1.2 Motivação

A motivação deste trabalho decorre das limitações observadas nos paradigmas diagnósticos atualmente empregados na identificação de transtornos de ansiedade em crianças, especialmente em idade pré-escolar. A predominância de abordagens clínicas baseadas em instrumentos subjetivos e em classificações dicotômicas compromete a detecção precoce e a estratificação adequada do risco, dificultando a definição de intervenções proporcionais à gravidade do quadro clínico.

Do ponto de vista clínico, a ausência de mecanismos capazes de discriminar níveis graduais de risco pode resultar em intervenções tardias, subtratamento de casos mais severos ou alocação ineficiente de recursos terapêuticos, conforme argumentado por Kazdin [24] e Craske et al. [25]. Em populações pediátricas, essas limitações assumem maior relevância, uma vez que atrasos na identificação de quadros ansiosos estão associados a prejuízos persistentes no desenvolvimento emocional, cognitivo e social, com repercussões ao longo do ciclo vital, conforme discutido por Smith et al. [26] e Dickson et al. [27].

Sob a perspectiva da Engenharia Elétrica, esse cenário pode ser interpretado como um problema de processamento e análise de sinais complexos, no qual variáveis psicométricas e

comportamentais representam observações ruidosas de estados emocionais latentes. A modelagem computacional desses dados requer técnicas capazes de capturar relações não lineares, interações multivariadas e padrões estatísticos sutis, reduzindo a subjetividade inerente ao julgamento humano e ampliando a consistência das decisões clínicas.

Nesse contexto, métodos de aprendizado de máquina supervisionado emergem como ferramentas adequadas para apoiar a classificação do risco de ansiedade pediátrica de forma objetiva, reprodutível e escalável. A motivação central deste trabalho reside, portanto, na necessidade de investigar modelos computacionais capazes de realizar classificações multiníveis do risco clínico, alinhadas à natureza dimensional dos transtornos ansiosos e fundamentadas em princípios da Engenharia Elétrica e do processamento estatístico de sinais.

### 1.3 Objetivos

Esta seção apresenta o objetivo geral e os objetivos específicos que orientam o desenvolvimento desta pesquisa.

#### 1.3.1 Objetivo Geral

Desenvolver e avaliar um sistema computacional de apoio à decisão para a classificação binária e multinível do risco de transtornos de ansiedade em crianças em idade pré-escolar.

#### 1.3.2 Objetivos Específicos

- Implementar e comparar modelos de aprendizado de máquina supervisionado amplamente consolidados na literatura, na classificação do risco de ansiedade pediátrica em cenários binários e multiníveis;
- Avaliar o desempenho e a capacidade de generalização dos modelos por meio de métricas quantitativas padronizadas, considerando critérios de relevância clínica;
- Investigar a coerência das predições dos modelos em relação aos referenciais diagnósticos internacionais estabelecidos pelo DSM-5 (*American Psychiatric Association* [10]) e pela CID-11 (*World Health Organization* [11]);
- Propor um *pipeline* metodológico reprodutível, fundamentado em princípios da Engenharia Elétrica, para aplicação em sistemas computacionais de apoio à decisão clínica.

## **1.4 Organização do Trabalho**

Este trabalho está organizado da seguinte forma. O Capítulo 2 apresenta o referencial teórico que fundamenta a pesquisa, contemplando conceitos de Engenharia Elétrica, aprendizado de máquina e saúde mental infantil. O Capítulo 3 descreve a metodologia proposta, incluindo a caracterização da base de dados, as etapas de pré-processamento, os modelos implementados e o protocolo de validação adotado. O Capítulo 4 apresenta os resultados experimentais e a discussão dos achados à luz dos objetivos propostos. Por fim, o Capítulo 5 expõe as conclusões, limitações do estudo e perspectivas para trabalhos futuros.

## 2 REFERENCIAL TEÓRICO

### 2.1 Transtorno de Ansiedade Pediátrica

O Transtorno de Ansiedade constitui um conjunto de condições psicopatológicas caracterizadas pela presença persistente e excessiva de medo, preocupação antecipatória e respostas de hipervigilância, que se manifestam de forma desproporcional em relação às demandas reais do ambiente e comprometem o funcionamento emocional, cognitivo e comportamental do indivíduo. Diferentemente de respostas adaptativas normativas ao estresse, a ansiedade patológica apresenta intensidade, duração e impacto funcional elevados, interferindo de maneira significativa na regulação emocional e na tomada de decisão, conforme descrito por Barlow [28].

Sob a perspectiva clínica contemporânea, os transtornos de ansiedade englobam múltiplas categorias diagnósticas, incluindo transtorno de ansiedade generalizada, ansiedade de separação, fobias específicas, ansiedade social e transtorno do pânico, entre outros. Embora esses quadros apresentem critérios diagnósticos distintos, compartilham mecanismos subjacentes relacionados à avaliação distorcida de ameaça, à antecipação negativa de eventos futuros e à ativação persistente de sistemas neurobiológicos associados ao medo, conforme discutido por LeDoux [29].

Do ponto de vista neurofuncional, a ansiedade está associada à disfunção de circuitos cerebrais responsáveis pela detecção de ameaça, regulação emocional e controle cognitivo, envolvendo estruturas como a amígdala, o hipocampo e regiões do córtex pré-frontal. A hiperatividade da amígdala, combinada a mecanismos insuficientes de modulação cortical, resulta em respostas emocionais amplificadas e dificuldade de extinção do medo, configurando um estado de vigilância persistente, conforme descrito por Davis [30].

Em termos de Engenharia Elétrica aplicada à saúde, o Transtorno de Ansiedade pode ser interpretado como uma alteração no comportamento dinâmico de um sistema complexo de regulação emocional, no qual mecanismos de realimentação e controle tornam-se desbalanceados. Nesse enquadramento, estímulos ambientais são processados por sistemas neurais que operam como filtros adaptativos, responsáveis por discriminar sinais relevantes de ruído. Na ansiedade patológica, observa-se uma redução da seletividade desses filtros, resultando em respostas exacerbadas a estímulos de baixa relevância ou ambiguidade, conforme argumentado por Thayer e Lane [31].

Além disso, o transtorno apresenta natureza multidimensional, envolvendo componentes

afetivos, cognitivos, comportamentais e fisiológicos. Esses componentes não se manifestam de forma isolada, mas interagem de maneira não linear, produzindo padrões complexos de resposta que variam entre indivíduos e ao longo do tempo. Essa característica dificulta a representação do transtorno por modelos determinísticos simples, exigindo abordagens que considerem variabilidade, incerteza e interdependência entre múltiplas variáveis observáveis, conforme discutido por Kendall [32].

No contexto pediátrico, essas características tornam-se ainda mais pronunciadas. Crianças frequentemente apresentam manifestações ansiosas atípicas ou inespecíficas, como irritabilidade, queixas somáticas e alterações comportamentais, em vez de verbalizações explícitas de medo ou preocupação. Essa apresentação clínica reforça a compreensão do Transtorno de Ansiedade como um fenômeno latente, cuja identificação depende da inferência a partir de múltiplos indicadores indiretos, conforme demonstrado por Costello et al. [33].

Assim, o Transtorno de Ansiedade pode ser conceituado, sob uma perspectiva sistêmica, como um estado latente contínuo associado à disfunção dos mecanismos de avaliação de ameaça e regulação emocional, observável apenas de forma indireta por meio de sinais psicométricos, comportamentais e fisiológicos. Essa definição estabelece a base conceitual para a formulação do problema de inferência e estratificação de risco em saúde mental pediátrica, fundamento que dialoga diretamente com a modelagem multinível proposta neste trabalho.

Em termos epidemiológicos e clínicos, sintomas ansiosos apresentam prevalência relevante em populações pediátricas, com início frequentemente precoce e tendência à persistência ao longo do desenvolvimento quando não identificados e manejados adequadamente, conforme evidenciado por Costello et al. [33] e Egger et al. [34]. Sob a perspectiva sistêmica, tal dinâmica evolutiva reforça a necessidade de modelagens quantitativas sob incerteza, compatíveis com a natureza contínua, multifatorial e ordinal do risco clínico.

Em crianças, os quadros ansiosos manifestam-se por meio de padrões heterogêneos de resposta emocional, comportamental e fisiológica, incluindo medo excessivo, evitação, hiper-vigilância e alterações na regulação autonômica. Essas manifestações apresentam elevada variabilidade interindividual e frequentemente se sobrepõem a sintomas de outros transtornos do neurodesenvolvimento, como TDAH e transtornos do humor, dificultando a definição de fronteiras diagnósticas bem delimitadas, conforme discutido por Angold et al. [35].

Do ponto de vista sistêmico, a ansiedade pediátrica não pode ser compreendida como um evento pontual ou binário, mas como um processo dinâmico, contínuo e multifatorial, no qual fa-

tores neurobiológicos, ambientais, cognitivos e comportamentais interagem de forma não linear ao longo do tempo. Essa característica afasta o problema de formulações determinísticas simples e aproxima-o de paradigmas clássicos de inferência estatística sob incerteza, amplamente estudados na Engenharia Elétrica e alinhados à modelagem ordinal adotada nesta pesquisa.

### 2.1.1 Referenciais Diagnósticos: DSM-5 e CID-11

A caracterização clínica dos transtornos de ansiedade, bem como a definição formal dos critérios utilizados para identificação, classificação e avaliação da gravidade dos sintomas, é estabelecida por sistemas diagnósticos padronizados amplamente adotados na prática clínica e na pesquisa em saúde mental. Entre esses sistemas, destacam-se o *Manual Diagnóstico e Estatístico de Transtornos Mentais – Quinta Edição (Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition)*, publicado pela American Psychiatric Association (APA), e a *Classificação Internacional de Doenças – 11ª Revisão (International Classification of Diseases)*, desenvolvida pela *World Health Organization (WHO)* [10, 11].

O DSM-5 estabelece categorias diagnósticas baseadas em conjuntos de critérios clínicos que incluem a presença, a duração e a intensidade dos sintomas, bem como o grau de prejuízo funcional associado. No contexto dos transtornos de ansiedade, esse manual descreve quadros específicos — como transtorno de ansiedade generalizada, ansiedade de separação, ansiedade social e fobias específicas — enfatizando a distinção entre respostas adaptativas ao estresse e padrões persistentes de ansiedade patológica. Tal distinção fundamenta-se na combinação entre sofrimento subjetivo clinicamente significativo e impacto funcional mensurável nos domínios social, acadêmico e familiar.

De forma complementar, a CID-11 constitui um sistema diagnóstico de abrangência internacional, com foco na padronização terminológica e na aplicabilidade em diferentes contextos culturais e sistemas de saúde. A CID-11 adota uma estrutura conceitual amplamente alinhada ao DSM-5 no que se refere à caracterização dos transtornos de ansiedade, incorporando critérios baseados na persistência dos sintomas, na desproporcionalidade das respostas emocionais e no prejuízo funcional associado. A convergência entre esses referenciais reforça a compreensão da ansiedade patológica como um fenômeno clínico de natureza dimensional.

Sob a perspectiva da Engenharia Elétrica aplicada à saúde, o DSM-5 e a CID-11 podem ser interpretados como modelos conceituais normativos que delimitam regiões do espaço de estados clínicos correspondentes a diferentes níveis de comprometimento emocional e fun-

cional. Entretanto, tais sistemas operam sobre informações inferidas indiretamente a partir de instrumentos psicométricos, entrevistas clínicas e observações comportamentais, os quais estão sujeitos a ruído, viés e variabilidade interindividual. Assim, mesmo na presença de critérios diagnósticos formalmente definidos, a atribuição de rótulos clínicos envolve incerteza intrínseca, impactando diretamente a separabilidade entre classes e a confiabilidade do *ground truth* utilizado em modelos supervisionados — aspecto que dialoga com os resultados observados nas classes intermediárias deste estudo.

Além disso, ambos os sistemas reconhecem implicitamente a existência de estados intermediários entre a ausência de transtorno e o diagnóstico plenamente estabelecido, particularmente em populações pediátricas. A exigência de limiares de duração, intensidade e prejuízo funcional implica que crianças podem apresentar sintomas clinicamente relevantes sem satisfazer todos os critérios categóricos formais. Esse aspecto reforça a necessidade de abordagens baseadas em estratificação de risco, nas quais o objetivo consiste em estimar graus progressivos de vulnerabilidade, em vez de impor decisões binárias rígidas sobre um fenômeno de natureza contínua, dinâmica e multifatorial.

### **2.1.2 Transtorno de Ansiedade Generalizada (GAD) e Ansiedade de Separação (SAD)**

Entre os transtornos de ansiedade mais relevantes na infância e no contexto do presente trabalho, destacam-se o Transtorno de Ansiedade Generalizada (*Generalized Anxiety Disorder – GAD*) e o Transtorno de Ansiedade de Separação (*Separation Anxiety Disorder – SAD*), ambos descritos no DSM-5 [10] e na CID-11 [11].

O *GAD* caracteriza-se pela presença de preocupação excessiva e persistente, de difícil controle, associada a sintomas somáticos e cognitivos como inquietação, tensão muscular, fadiga, irritabilidade e dificuldade de concentração. Em populações pediátricas, essas manifestações podem apresentar-se de forma menos verbalizada e mais comportamental, incluindo necessidade frequente de *reassurance*, evitamento e queixas físicas recorrentes, reforçando sua natureza parcialmente observável por indicadores indiretos.

O *SAD*, por sua vez, refere-se a um padrão de medo ou ansiedade excessiva diante da separação de figuras de apego, com manifestações como sofrimento intenso antecipatório, recusa escolar, preocupação persistente com eventos de perda e sintomas físicos diante da separação. Em idade pré-escolar, o desafio clínico reside em distinguir respostas normativas do desenvolvimento de quadros persistentes e desproporcionais, cuja intensidade e prejuízo funcional sugerem

risco clínico aumentado.

No escopo deste estudo, a explicitação conceitual de *GAD* e *SAD* é particularmente relevante, uma vez que o processo de modelagem busca inferir, a partir de variáveis psicométricas tabulares, padrões compatíveis com níveis progressivos de risco associados a tais construtos diagnósticos, aproximando a modelagem computacional da lógica clínica dimensional.

### **2.1.3 Dificuldades no Diagnóstico Precoce**

O diagnóstico precoce da ansiedade pediátrica apresenta limitações estruturais que podem ser formalizadas, sob a perspectiva da Engenharia Elétrica, como um problema de observação indireta de estados latentes. Os estados emocionais associados à ansiedade não são diretamente mensuráveis, sendo inferidos a partir de variáveis observáveis ruidosas, tais como respostas a questionários, escalas psicométricas e registros comportamentais.

Essas observações apresentam natureza subjetiva e dependem fortemente do informante, seja ele a própria criança, pais ou professores. De Los Reyes e Kazdin [36] demonstraram baixa concordância entre diferentes fontes de informação, indicando que o vetor de observações disponível para o diagnóstico está sujeito a variabilidade sistemática e ruído não gaussiano. Sob a ótica da teoria de sistemas, isso implica que o problema de inferência opera em regime de observabilidade parcial.

Adicionalmente, limitações cognitivas e linguísticas inerentes às faixas etárias mais jovens restringem a confiabilidade de autorrelatos, aumentando a dependência de medidas indiretas. Conforme discutido por Del Carmen-Wiggins e Carter [37], esse cenário amplia a incerteza associada à estimação do estado emocional real da criança, favorecendo atrasos diagnósticos e elevando a probabilidade de decisões subótimas em contextos clínicos.

Sob essa perspectiva, a modelagem computacional proposta neste trabalho pode ser compreendida como uma estratégia de redução de incerteza estatística, buscando estimar níveis graduais de risco a partir de múltiplos indicadores observáveis, respeitando a natureza dimensional, contínua e parcialmente observável do construto ansiedade pediátrica.

### **2.1.4 Necessidade de Estratificação de Risco como Problema de Decisão**

Diante das limitações do diagnóstico categórico precoce, a literatura tem apontado a estratificação de risco como uma abordagem conceitualmente mais adequada para a ansiedade pediátrica, conforme discutido por Pine et al. [38]. Em termos de Engenharia Elétrica, essa

abordagem pode ser interpretada como um problema de decisão multinível, no qual o objetivo não é apenas classificar estados extremos, mas estimar graus progressivos de risco associados a trajetórias clínicas distintas.

A ansiedade infantil frequentemente emerge em níveis subclínicos e evolui gradualmente, caracterizando um processo de transição contínua entre estados de baixa, moderada e alta vulnerabilidade. A adoção de esquemas binários de decisão ignora essa dinâmica e impõe uma quantização excessivamente grosseira sobre um fenômeno intrinsecamente contínuo, resultando em perda de informação relevante para a tomada de decisão clínica, conforme argumentado por Kraemer et al. [39].

Sob essa perspectiva, a estratificação de risco pode ser formalizada como a estimação de um estado ordinal latente, no qual cada nível representa uma região do espaço de estados associada a diferentes probabilidades de progressão, agravamento ou remissão dos sintomas. Essa formulação aproxima o problema de paradigmas clássicos de classificação ordinal e decisão sob incerteza, amplamente explorados na Engenharia Elétrica e na teoria da detecção.

## 2.2 Psicometria como Sinal Discreto Ruidoso

Os dados disponíveis para a análise da ansiedade pediátrica consistem predominantemente em variáveis psicométricas e comportamentais estruturadas em formato tabular. Sob a ótica da Engenharia Elétrica, tais variáveis podem ser interpretadas como amostras discretizadas e quantizadas de processos psicológicos latentes contínuos.

Seja  $s(t)$  um estado emocional latente associado ao nível de ansiedade de uma criança ao longo do tempo. As medidas psicométricas disponíveis correspondem a observações discretas desse estado, sujeitas a ruído e distorções introduzidas por fatores cognitivos, contextuais e instrumentais. Esse processo pode ser modelado de forma abstrata como:

$$x[n] = Q\{s(t_n) + \varepsilon[n]\}, \quad (1)$$

em que  $Q\{\cdot\}$  representa o operador de quantização associado à escala psicométrica utilizada,  $t_n$  o instante de amostragem e  $\varepsilon[n]$  um termo de ruído que incorpora variabilidade subjetiva e erro de medição.

Essa formulação evidencia que o problema não consiste em classificar sinais ideais, mas em inferir estados latentes a partir de observações ruidosas, incompletas e altamente correla-

cionadas. Como consequência, a separabilidade entre diferentes níveis de risco é limitada por fatores estruturais do próprio processo de medição, e não apenas pela escolha do modelo computacional, conforme discutido por Johnston e Pennypacker [40]. Esse enquadramento fundamenta a adoção de modelos supervisionados não lineares e justifica a necessidade de estratégias de avaliação compatíveis com ordinalidade e assimetria de custo do erro, conforme operacionalizado no pipeline metodológico deste trabalho.

### **2.2.1 Subjetividade, Ruído e Variabilidade Interindividual**

A presença de subjetividade constitui uma característica estrutural dos dados psicométricos, e não uma limitação circunstancial. Respostas a instrumentos de avaliação refletem não apenas o estado emocional da criança, mas também expectativas do informante, contexto socio-cultural e variações temporais do comportamento, conforme analisado por Weems et al. [41].

Do ponto de vista estatístico, essa subjetividade pode ser interpretada como uma combinação de ruído aleatório e viés sistemático, resultando em distribuições de classes parcialmente sobrepostas. Em termos de Engenharia Elétrica, isso implica que o problema apresenta uma relação sinal-ruído limitada, impondo um limite superior teórico ao desempenho de qualquer sistema de classificação.

Adicionalmente, a variabilidade interindividual no desenvolvimento neuropsicológico introduz heterogeneidade estrutural no processo gerador dos dados, dificultando a definição de modelos universais e reforçando a necessidade de abordagens capazes de lidar com heterogeneidade populacional.

### **2.2.2 Natureza Ordinal e Multinível do Risco Clínico**

A ansiedade pediátrica apresenta natureza essencialmente ordinal, na qual os níveis de risco representam graus progressivos de severidade clínica e impacto funcional. Diferentemente de problemas de classificação nominal, as classes associadas ao risco de ansiedade possuem uma relação de ordem semântica, de modo que erros de classificação apresentam consequências clínicas assimétricas, conforme discutido por Kraemer et al. [39].

A confusão entre níveis adjacentes de risco tende a ser menos crítica do que deslocamentos amplos no espectro ordinal. Essa assimetria indica que a avaliação do problema não deve se restringir exclusivamente a métricas de acerto global, exigindo uma formulação que considere explicitamente a magnitude do erro de decisão. Sob a ótica da Engenharia Elétrica, essa carac-

terística aproxima o problema de sistemas de decisão com custos de erro não uniformes, nos quais a função de perda deve refletir não apenas a ocorrência do erro, mas também sua gravidade relativa. Essa compreensão fundamenta a necessidade de modelos e métricas compatíveis com a natureza ordinal do risco clínico, estabelecendo as bases conceituais para as abordagens teóricas de solução discutidas nas seções subsequentes.

## 2.3 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNAs) constituem uma classe de modelos computacionais inspirados nos princípios estruturais e funcionais dos sistemas neurais biológicos, sendo amplamente empregadas na modelagem de sistemas complexos caracterizados por relações não lineares e alta dimensionalidade. Sob a perspectiva da Engenharia Elétrica, as RNAs podem ser interpretadas como sistemas de processamento de sinais distribuídos, nos quais informações de entrada são transformadas por meio de operações matemáticas paramétricas ajustáveis, resultando em representações internas capazes de capturar padrões latentes nos dados.

Formalmente, uma RNA é composta por unidades elementares denominadas neurônios artificiais, interconectadas por pesos sinápticos que modulam a contribuição relativa de cada sinal de entrada. O modelo seminal do perceptron foi originalmente proposto por Rosenblatt [42], e sua formalização moderna é amplamente descrita por Goodfellow et al. [43]. Cada neurônio executa uma operação matemática que consiste na soma ponderada das entradas, acrescida de um termo de viés, seguida da aplicação de uma função de ativação não linear:

$$a_j = \phi \left( \sum_{i=1}^n w_{ij} x_i + b_j \right) \quad (2)$$

em que  $a_j$  denota a saída (ativação) do neurônio  $j$ ,  $w_{ij}$  representa o peso associado à conexão entre o neurônio  $i$  e o neurônio  $j$ ,  $x_i$  corresponde ao sinal de entrada, e  $b_j$  é o termo de viés. A função de ativação  $\phi(\cdot)$  introduz não linearidade ao modelo, ampliando sua capacidade expressiva e permitindo a aproximação de funções complexas. Funções como a sigmoide, a tangente hiperbólica e a *Rectified Linear Unit (ReLU)* são amplamente utilizadas devido às suas propriedades matemáticas e computacionais.

No contexto do aprendizado supervisionado, o processo de treinamento consiste no ajuste iterativo dos pesos sinápticos com o objetivo de minimizar uma função de custo  $\mathcal{L}$ , que quantifica a discrepância entre as saídas previstas pela rede e os valores de referência. O algoritmo de

treinamento mais difundido é o *backpropagation*, formalizado por Rumelhart et al.[44], fundamentado no cálculo do gradiente da função de perda em relação aos parâmetros do modelo e na aplicação do método do gradiente descendente:

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial \mathcal{L}}{\partial w_{ij}} \quad (3)$$

em que  $\eta$  representa a taxa de aprendizado, responsável por controlar a magnitude das atualizações dos pesos ao longo das iterações. Essa expressão formaliza o mecanismo de correção de erro que orienta a convergência do modelo em direção a um mínimo local da função de custo.

Em tarefas de classificação, uma das funções de perda mais empregadas é a entropia cruzada, amplamente discutida por Goodfellow et al.[43], definida como:

$$\mathcal{L} = - \sum_{i=1}^C y_i \log(\hat{y}_i) \quad (4)$$

na qual  $y_i$  corresponde ao rótulo verdadeiro da classe  $i$  e  $\hat{y}_i$  denota a probabilidade estimada pela rede para essa classe. Essa função penaliza de forma mais severa predições incorretas associadas a altas probabilidades, favorecendo modelos com melhor calibração probabilística e maior capacidade discriminativa.

Dessa forma, as redes neurais artificiais constituem ferramentas matematicamente robustas para a modelagem de fenômenos complexos, permitindo a integração de múltiplas variáveis e a extração de padrões não triviais a partir de dados ruidosos. Tais propriedades justificam sua ampla adoção em aplicações de Engenharia Elétrica voltadas à análise de sinais, reconhecimento de padrões e apoio à decisão em sistemas complexos, incluindo o domínio da saúde mental pediátrica.

### 2.3.1 *Multilayer Perceptron (MLP)*

O *Multilayer Perceptron (MLP)* constitui uma das arquiteturas mais consolidadas de redes neurais artificiais e desempenha papel fundamental no desenvolvimento histórico dos métodos de aprendizado supervisionado. Trata-se de um modelo do tipo *feedforward*, no qual o fluxo de informação ocorre de maneira unidirecional, da camada de entrada em direção à camada de saída, atravessando uma ou mais camadas ocultas intermediárias. Cada camada realiza transformações lineares seguidas pela aplicação de funções de ativação não lineares, conferindo à rede

elevada capacidade de aproximação funcional e permitindo a modelagem de relações complexas e não lineares entre variáveis. O teorema da aproximação universal, formalizado por Hornik et al. [15], estabelece que uma *MLP* com ao menos uma camada oculta e funções de ativação não lineares adequadas pode aproximar qualquer função contínua em um conjunto compacto, sob determinadas condições.

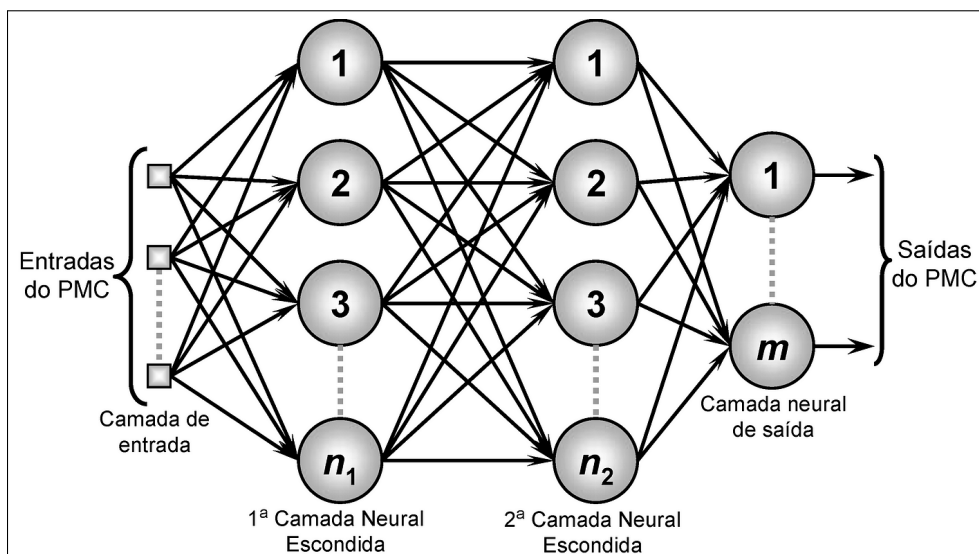
O processo de aprendizado do *MLP* é conduzido de forma supervisionada por meio do algoritmo de retropropagação do erro (*backpropagation*), formalizado por Rumelhart et al. [44], que possibilita o ajuste iterativo dos pesos sinápticos a partir de exemplos rotulados. Sob a perspectiva da Engenharia Elétrica, o *MLP* pode ser interpretado como um sistema não linear parametrizado, no qual sinais de entrada são progressivamente transformados em representações internas de maior nível de abstração. Matematicamente, as operações associadas a uma *MLP* com uma camada oculta podem ser descritas pelas seguintes expressões:

$$\mathbf{h} = \phi (\mathbf{W}^{(1)}\mathbf{x} + \mathbf{b}^{(1)}) \quad (5)$$

$$\hat{\mathbf{y}} = \psi (\mathbf{W}^{(2)}\mathbf{h} + \mathbf{b}^{(2)}) \quad (6)$$

em que  $\mathbf{x}$  representa o vetor de entrada,  $\mathbf{W}^{(1)}$  e  $\mathbf{W}^{(2)}$  correspondem às matrizes de pesos associadas às camadas oculta e de saída, respectivamente,  $\mathbf{b}^{(1)}$  e  $\mathbf{b}^{(2)}$  são os vetores de vies,  $\phi(\cdot)$  e  $\psi(\cdot)$  denotam funções de ativação — como *Rectified Linear Unit (ReLU)*, *sigmoid* ou *softmax* — e  $\hat{\mathbf{y}}$  representa o vetor de saída da rede.

**Figura 1 – Representação conceitual da arquitetura de um *Multilayer Perceptron (MLP)***



Fonte: Representação conceitual adaptada pela autora a partir de Silva, Spatti e Flauzino [45].

As *MLPs* têm sido amplamente empregadas em aplicações de predição clínica e biomédica devido à sua capacidade de modelar interações não lineares entre variáveis psicométricas, comportamentais e psicofisiológicas, características comuns em dados provenientes de contextos clínicos reais, conforme demonstrado por Strauss et al. [46] e Zhang et al. [47]. No presente estudo, a *MLP* é adotada como arquitetura supervisionada de referência, permitindo a avaliação de seu desempenho preditivo e de sua capacidade de generalização em comparação com outros modelos investigados, como *Random Forest (RF)* e *Support Vector Machine (SVM)*.

O processo de treinamento da *MLP* envolve a minimização iterativa da função de custo, sendo frequentemente conduzido por meio de otimizadores baseados em gradiente, como o *Adaptive Moment Estimation (Adam)*, proposto por Diederik Kingma e Jimmy Ba [48]. Esse método combina estimativas adaptativas dos momentos de primeira e segunda ordem do gradiente, ajustando dinamicamente a taxa de aprendizado e promovendo convergência mais estável e eficiente, mesmo em espaços de parâmetros altamente não convexos.

### 2.3.2 *Support Vector Machines (SVM)*

As *Support Vector Machines (SVM)* constituem uma classe de algoritmos de aprendizado de máquina supervisionado amplamente empregada em problemas de classificação e regressão, destacando-se por sua fundamentação geométrica e por propriedades robustas de generalização. O método foi originalmente proposto por Corinna Cortes e Vladimir Vapnik [12] e posteriormente formalizado em maior profundidade por Vladimir Vapnik no contexto da teoria do aprendizado estatístico. Sob a perspectiva da Engenharia Elétrica, as *SVMs* podem ser interpretadas como métodos de otimização convexa aplicados à separação de sinais em espaços de características, nos quais o objetivo consiste em determinar uma fronteira de decisão que maximize a margem entre diferentes classes.

No caso linearmente separável, o problema central da *SVM* consiste em encontrar um hiperplano definido por um vetor normal  $\mathbf{w}$  e um termo de viés  $b$ , tal que:

$$\mathbf{w} \cdot \mathbf{x} + b = 0 \quad (7)$$

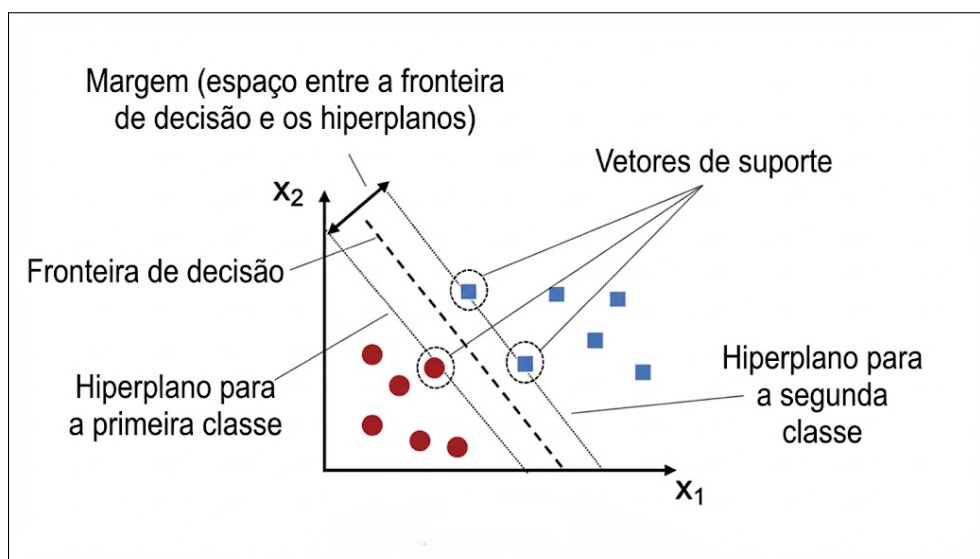
Esse hiperplano atua como fronteira de decisão, separando os dados de entrada  $\mathbf{x}$  em diferentes classes. A margem geométrica é definida como a distância entre o hiperplano de decisão e os hiperplanos paralelos que passam pelos vetores de suporte, sendo maximizada durante o processo de treinamento. Formalmente, o problema de otimização pode ser expresso como:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{sujeito a} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \quad \forall i \quad (8)$$

em que  $(\mathbf{x}_i, y_i)$  representam os pares de dados de entrada e rótulos associados. Essa formulação evidencia o compromisso entre separação correta das classes e maximização da margem, propriedade diretamente relacionada à capacidade de generalização do modelo.

Para cenários em que os dados não são linearmente separáveis, introduzem-se variáveis de folga e um parâmetro de regularização  $C$ , resultando no problema de margem suave (*soft margin*). Além disso, por meio do chamado *kernel trick*, conceito formalizado na teoria de máquinas de kernel por Schölkopf e Smola [49], as *SVMs* permitem mapear os dados de entrada para espaços de maior dimensionalidade, nos quais a separação linear torna-se viável. Funções *kernel* amplamente utilizadas incluem os *kernels* linear, polinomial e de base radial (*Radial Basis Function – RBF*), sendo este último particularmente eficaz na modelagem de fronteiras de decisão não lineares.

**Figura 2 – Representação conceitual do princípio de separação por *Support Vector Machines (SVM)***



**Fonte:** Representação conceitual adaptada pela autora a partir de [50].

Na Figura 2, ilustra-se o conceito geométrico fundamental das *SVMs*, destacando-se a fronteira de decisão, os hiperplanos de margem máxima e os vetores de suporte, que correspondem às amostras mais próximas da fronteira e exercem influência direta na definição do classificador. Esse caráter geométrico confere às *SVMs* elevada robustez frente a ruídos e *outliers* moderados, característica relevante em bases clínicas compostas por dados psicométricos e comportamentais heterogêneos.

No contexto da saúde mental pediátrica, as *Support Vector Machines (SVM)* têm sido

amplamente empregadas na classificação de estados emocionais e transtornos do neurodesenvolvimento, em virtude de sua capacidade de lidar com conjuntos de dados de dimensionalidade moderada e tamanho amostral limitado, mantendo bom desempenho preditivo, conforme discutido por Shatte et al.[6]. No presente trabalho, a *SVM* é utilizada como um dos modelos supervisionados de referência, permitindo a comparação de seu desempenho com outras abordagens de aprendizado de máquina, como *Random Forest (RF)* e *Multilayer Perceptron (MLP)*, na tarefa de classificação binária e multinível do risco de ansiedade pediátrica.

### 2.3.3 *Random Forest (RF)*

O *Random Forest (RF)* constitui um algoritmo supervisionado baseado em *ensembles* de árvores de decisão, originalmente proposto por Breiman [13], sendo amplamente empregado em tarefas de classificação e regressão devido à sua robustez estatística, elevada capacidade de generalização e desempenho consistente em cenários com variáveis heterogêneas e distribuições não gaussianas. Sob a ótica da Engenharia Elétrica, o *Random Forest* pode ser interpretado como um sistema de decisão composto, no qual múltiplos classificadores operam em paralelo sobre subconjuntos estocásticos dos dados e das variáveis de entrada, produzindo uma inferência agregada menos sensível a ruídos e variações amostrais.

O princípio central do *Random Forest* fundamenta-se no *bagging (bootstrap aggregating)*, técnica também introduzida por Breiman, na qual cada árvore é treinada a partir de uma amostra obtida via reamostragem com reposição do conjunto original. Adicionalmente, em cada nó de decisão, a seleção do atributo candidato à divisão é realizada a partir de um subconjunto aleatório de variáveis. Essa dupla aleatoriedade — sobre amostras e atributos — reduz a correlação entre árvores individuais e promove redução de variância, resultando em melhor generalização quando comparado a uma única árvore de decisão.

Em problemas de classificação, a predição do *Random Forest* pode ser formalizada como uma votação majoritária entre  $T$  árvores de decisão:

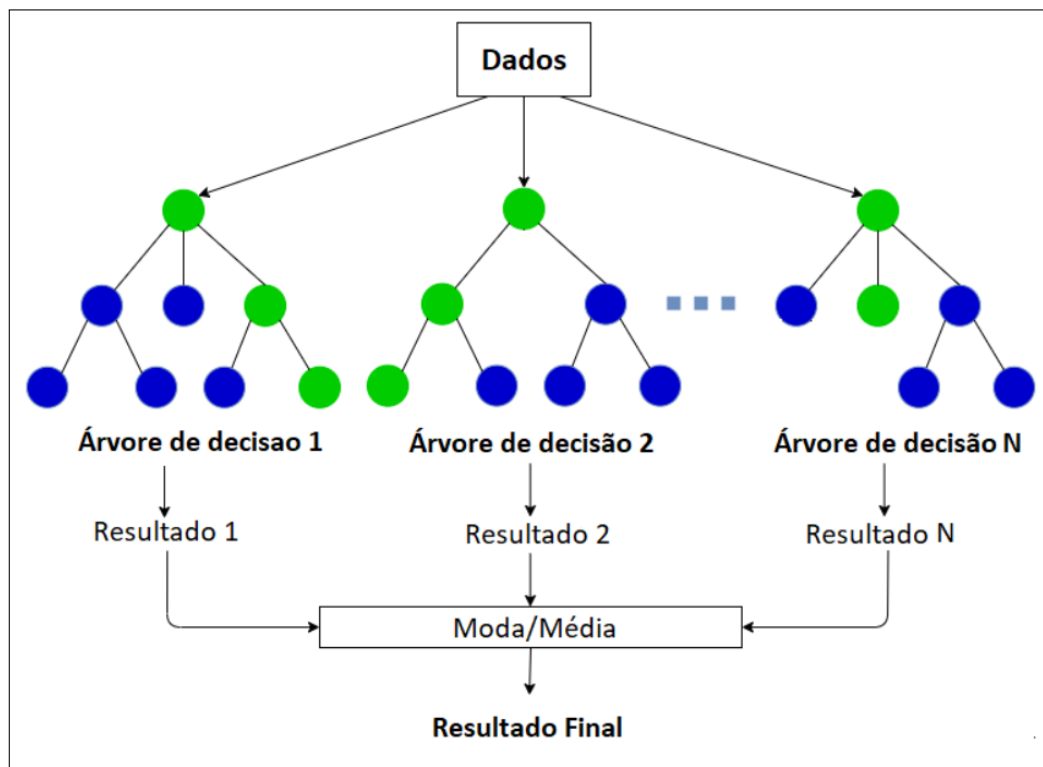
$$\hat{y} = \text{mode} \{h_t(\mathbf{x})\}_{t=1}^T, \quad (9)$$

em que  $h_t(\mathbf{x})$  denota a classe predita pela  $t$ -ésima árvore para a amostra  $\mathbf{x}$ . Em tarefas de regressão, a agregação é tipicamente realizada por meio da média aritmética das saídas individuais. Essa estratégia de combinação confere ao modelo elevada estabilidade, sendo particularmente adequada para dados clínicos e psicométricos, nos quais a variabilidade interindividual e

o ruído de medição são fatores inerentes.

Cada árvore que compõe o *Random Forest* é construída por divisões hierárquicas guiadas por critérios de impureza, como o índice de Gini ou a entropia, buscando maximizar o ganho informacional a cada partição. Embora árvores individuais sejam modelos de alta variância, o mecanismo de agregação do *Random Forest* mitiga esse efeito, preservando a capacidade de capturar relações não lineares e interações complexas entre variáveis sem incorrer, de maneira pronunciada, em sobreajuste.

**Figura 3 – Representação conceitual da estrutura de um *Random Forest* (RF)**



Fonte: Representação conceitual adaptada pela autora a partir de [51].

A Figura 3 ilustra o princípio de funcionamento do *Random Forest*: múltiplas árvores de decisão são induzidas a partir de reamostragens do conjunto de dados, produzindo saídas parciais que são então agregadas por votação majoritária (classificação) ou média (regressão) para compor a decisão final.

No contexto da saúde mental pediátrica, o *Random Forest* (RF) é frequentemente adotado por sua capacidade de lidar com conjuntos de dados tabulares de dimensão moderada, por tolerar atributos correlacionados e por fornecer estimativas de importância de variáveis, o que favorece interpretações quantitativas sobre quais atributos contribuem mais fortemente para a discriminação entre classes de risco. No presente trabalho, o *Random Forest* (RF) é utilizado como um dos modelos supervisionados de referência para a classificação binária e multinível

do risco de ansiedade pediátrica, permitindo a comparação de desempenho frente a outras abordagens supervisionadas empregadas no estudo.

### 2.3.4 Comparação entre Arquiteturas e Aspectos Computacionais

As arquiteturas investigadas neste referencial teórico — *MLP*, *SVM* e *RF* — representam paradigmas computacionais distintos aplicados à modelagem e classificação de fenômenos psicométricos e comportamentais. Tais abordagens diferenciam-se fundamentalmente quanto à natureza do aprendizado, à formulação matemática subjacente, ao grau de interpretabilidade e à capacidade de generalização frente à complexidade e à variabilidade inerentes aos dados clínicos. A Figura 4 sintetiza os principais aspectos comparativos entre essas arquiteturas, considerando parâmetros relevantes sob a ótica da Engenharia Elétrica aplicada à saúde mental.

**Figura 4 – Quadro comparativo das arquiteturas utilizadas no estudo**

Arquitetura	Tipo de Aprendizado	Vantagens	Limitações
<b>MLP</b>	<b>Supervisionado</b>	<ul style="list-style-type: none"> <li>▪ Modelagem de relações não lineares;</li> <li>▪ Elevada capacidade de aproximação funcional;</li> <li>▪ Adequado a dados tabulares psicométricos;</li> <li>▪ Integração com métodos de interpretabilidade.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Dependência de dados rotulados;</li> <li>▪ Sensibilidade à escolha de hiperparâmetros;</li> <li>▪ Menor transparência estrutural.</li> </ul>
<b>SVM</b>	<b>Supervisionado</b>	<ul style="list-style-type: none"> <li>▪ Formulação convexa; Maximização da margem;</li> <li>▪ Bom desempenho em bases de tamanho moderado;</li> <li>▪ Robustez frente a ruído e alta dimensionalidade.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Sensível à escolha do kernel e parâmetros;</li> <li>▪ Interpretabilidade limitada;</li> <li>▪ Custo computacional elevado em bases extensas.</li> </ul>
<b>Random Forest</b>	<b>Supervisionado (Ensemble)</b>	<ul style="list-style-type: none"> <li>▪ Alta robustez estatística; Boa generalização;</li> <li>▪ Tolerância a variáveis correlacionadas;</li> <li>▪ Estimativa de importância das variáveis.</li> </ul>	<ul style="list-style-type: none"> <li>▪ Modelo menos transparente globalmente;</li> <li>▪ Decisões agregadas dificultam rastreabilidade individual.</li> </ul>

Fonte: Elaboração própria (2026).

A complementaridade entre essas arquiteturas é evidente. As *Support Vector Machines (SVM)*, propostas por Cortes e Vapnik [12], oferecem uma abordagem geométrica rigorosa, fundamentada na otimização convexa e na maximização da margem, sendo particularmente eficazes em cenários com dimensionalidade moderada e amostras limitadas. O *Random Forest (RF)*, desenvolvido por Breiman [13], explora o princípio de agregação de múltiplos classificadores, resultando em elevada robustez estatística e capacidade de generalização, características

desejáveis em bases clínicas heterogêneas e ruidosas.

O *Multilayer Perceptron (MLP)*, fundamentado no trabalho clássico de Hornik et al. [15], destaca-se como a arquitetura mais flexível entre as avaliadas, permitindo a modelagem de relações altamente não lineares entre variáveis psicométricas, comportamentais e fisiológicas. Sua estrutura totalmente conectada é particularmente adequada a dados tabulares, como aqueles empregados neste estudo, além de favorecer a aplicação de técnicas de interpretabilidade baseadas em atribuição de importância, como o *SHapley Additive exPlanations (SHAP)*.

A escolha conjunta dessas três arquiteturas fundamenta-se na busca por um equilíbrio entre desempenho preditivo, robustez estatística e interpretabilidade clínica. Sob a perspectiva da Engenharia Elétrica, tal combinação permite analisar o problema de classificação do risco de ansiedade pediátrica a partir de diferentes paradigmas de decisão — geométrico, estatístico e conexionista — ampliando a compreensão computacional do fenômeno e fortalecendo a confiabilidade dos resultados. Dessa forma, a comparação entre *MLP*, *SVM* e *RF* não apenas evidencia suas diferenças estruturais e computacionais, mas também reforça a relevância de abordagens complementares na construção de sistemas inteligentes de apoio à decisão aplicados à saúde mental pediátrica.

### 2.3.5 Métodos de Explicabilidade: *SHAP* e Interpretação de Modelos

Em aplicações clínicas e biomédicas, a interpretabilidade dos modelos computacionais não constitui apenas um atributo desejável, mas um requisito ético, metodológico e regulatório. Sistemas de apoio à decisão baseados em Inteligência Artificial devem fornecer mecanismos transparentes que permitam compreender, auditar e justificar suas previsões, sobretudo quando aplicados a populações vulneráveis, como crianças. Conforme discutido por Doshi-Velez e Kim, bem como por Lipton [52, 53], a interpretabilidade é elemento central para garantir confiabilidade, rastreabilidade e alinhamento clínico das decisões automatizadas.

Neste contexto, adotou-se o método *SHapley Additive exPlanations (SHAP)*, proposto por Lundberg e Lee [54], como estratégia de interpretação do modelo preditivo. Fundamentado na teoria dos jogos cooperativos originalmente formulada por Lloyd Shapley, o *SHAP* quantifica a contribuição marginal de cada variável de entrada para a saída do modelo, considerando diferentes combinações possíveis de atributos. Essa formulação permite decompor a previsão individual de forma aditiva, preservando propriedades desejáveis como simetria e consistência na atribuição das contribuições marginais.

Formalmente, a predição de um modelo pode ser expressa como:

$$f(\mathbf{x}) = f_0 + \sum_{i=1}^M \phi_i, \quad (10)$$

em que  $f(\mathbf{x})$  representa a saída do modelo para a amostra  $\mathbf{x}$ ,  $f_0$  corresponde ao valor esperado da predição,  $M$  é o número total de variáveis e  $\phi_i$  denota o valor de Shapley associado à  $i$ -ésima característica. Cada  $\phi_i$  representa a contribuição marginal média da variável  $i$  para a predição final, avaliada sobre o espaço de possíveis coalizões.

Do ponto de vista computacional, o *SHAP* pode ser interpretado como uma ferramenta de análise de sensibilidade local e global, permitindo avaliar tanto a influência individual de cada variável em uma instância específica quanto padrões gerais de importância ao longo do conjunto de dados. Essa característica é particularmente relevante em modelos não lineares, como redes neurais artificiais, nos quais a interpretação direta dos parâmetros não é trivial.

No presente estudo, o *SHAP* foi aplicado ao modelo supervisionado selecionado como de melhor desempenho no cenário multinível, conforme detalhado no capítulo de resultados, com o objetivo de identificar quais variáveis psicométricas e comportamentais exercem maior influência na classificação do risco de ansiedade pediátrica. Essa análise contribui para a transparência algorítmica e possibilita avaliar a coerência clínica das decisões do modelo, verificando o alinhamento dos atributos mais relevantes com construtos diagnósticos reconhecidos pelo *DSM-5* [10] e pela *CID-11* [11].

Dessa forma, a incorporação do *SHAP* integra desempenho preditivo e interpretação clínica, aproximando os modelos computacionais das exigências práticas da psiquiatria do desenvolvimento. Sob a perspectiva da Engenharia Elétrica aplicada à saúde, essa abordagem reforça a confiabilidade dos sistemas inteligentes de apoio à decisão, promovendo análises quantitativas transparentes, auditáveis e metodologicamente consistentes.

## 2.4 Métricas de Avaliação de Desempenho

A mensuração objetiva do desempenho de modelos de aprendizado de máquina constitui uma etapa indispensável para a validação científica, estatística e clínica de sistemas de classificação aplicados à saúde, conforme discutido por Powers [55]. No contexto da modelagem do risco de transtornos mentais na infância, as métricas de desempenho assumem papel ainda mais crítico, uma vez que erros de classificação podem implicar consequências clínicas, éticas e soci-

ais relevantes. Assim, a avaliação não se restringe à eficácia computacional, mas deve refletir a confiabilidade diagnóstica, a robustez estatística e a segurança preditiva dos modelos propostos.

Nesta seção, descrevem-se as principais métricas empregadas na avaliação dos classificadores desenvolvidos neste estudo, com ênfase em sua fundamentação matemática, interpretação estatística e relevância clínica, particularmente no cenário de triagem precoce do risco de ansiedade pediátrica.

### 2.4.1 Matriz de Confusão

A matriz de confusão constitui a estrutura fundamental para a avaliação de classificadores supervisionados, conforme sistematizado por Raschka [56], pois explicita a correspondência entre as classes verdadeiras e as classes previstas pelo modelo. Para problemas de classificação binária, a matriz pode ser formalmente representada como:

	Predito: Positivo	Predito: Negativo
Real: Positivo	$TP$	$FN$
Real: Negativo	$FP$	$TN$

em que:

- $TP$  (*True Positives*) corresponde aos casos corretamente identificados como positivos, como crianças com risco elevado de ansiedade adequadamente detectadas;
- $TN$  (*True Negatives*) representa os casos corretamente classificados como negativos;
- $FP$  (*False Positives*) indica indivíduos incorretamente classificados como positivos;
- $FN$  (*False Negatives*) refere-se a casos positivos não identificados pelo modelo.

A matriz de confusão fornece uma visão detalhada dos acertos e erros do classificador, permitindo distinguir entre diferentes tipos de erro e servindo como base para o cálculo das métricas derivadas de desempenho. Tal distinção é particularmente relevante em aplicações clínicas, nas quais diferentes erros possuem impactos assimétricos. Em contextos de triagem em saúde mental pediátrica, por exemplo, a ocorrência de *false negatives* pode resultar na não identificação de crianças em situação de risco, configurando um impacto ético e clínico potencialmente mais grave do que a ocorrência de *false positives*. Por outro lado, modelos com perfil mais conservador — caracterizados por maior especificidade e menor taxa de *false positives* — podem reduzir alarmes indevidos, aspecto discutido nos resultados deste trabalho.

**Figura 5 – Representação conceitual da Matriz de Confusão**

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Fonte: Adaptado pela autora a partir de [57].

A partir da matriz de confusão, é possível derivar métricas quantitativas que capturam diferentes aspectos do desempenho do classificador, tais como acurácia, precisão, sensibilidade (ou *recall*), especificidade e *F1-score*. Essas métricas permitem uma avaliação multifacetada do modelo, conciliando critérios estatísticos e considerações clínicas, sendo posteriormente complementadas, neste trabalho, pela análise do erro clínico ordinal (erros adjacentes e não adjacentes) no cenário multinível.

Para o cenário multiclasse, as métricas foram agregadas por média *macro* (*macro-averaged*), isto é, uma média não ponderada por classe, atribuindo o mesmo peso a cada nível de risco. Essa escolha é particularmente relevante em bases com desbalanceamento entre níveis, pois evita que o desempenho do classificador seja dominado pelas classes majoritárias e favorece uma avaliação mais equitativa entre os diferentes estratos de risco clínico.

#### 2.4.2 Acurácia (*Accuracy*)

A acurácia corresponde à proporção de classificações corretas em relação ao número total de amostras avaliadas, sendo uma das métricas mais utilizadas na avaliação global de modelos de classificação, conforme discutido por Johnson et al.[58]. Do ponto de vista matemático, a acurácia é definida como:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

em que *TP* e *TN* representam, respectivamente, os números de verdadeiros positivos e verdadeiros negativos, enquanto *FP* e *FN* denotam os falsos positivos e falsos negativos.

Apesar de sua simplicidade e ampla difusão, a acurácia apresenta limitações importantes, especialmente em cenários caracterizados por desbalanceamento entre as classes. Em bases

clínicas, como aquelas relacionadas à avaliação do risco de ansiedade pediátrica, é comum que a classe majoritária corresponda a indivíduos com baixo ou moderado risco, enquanto os casos de risco elevado constituem uma parcela menor da amostra. Nessas circunstâncias, modelos triviais podem alcançar valores elevados de acurácia simplesmente ao privilegiar a classe majoritária, sem necessariamente apresentar capacidade discriminativa adequada.

Dessa forma, a acurácia deve ser interpretada com cautela, sendo mais informativa quando as classes apresentam distribuição aproximadamente balanceada ou quando se busca uma visão global do desempenho do sistema. Em aplicações de saúde mental, essa métrica fornece um indicativo geral de eficiência preditiva, mas não é suficiente para caracterizar, de forma isolada, a utilidade clínica do modelo.

### 2.4.3 Precisão (*Precision*) e Sensibilidade (*Recall*)

As métricas de precisão e sensibilidade permitem analisar aspectos complementares do desempenho de classificadores, especialmente relevantes em cenários clínicos nos quais diferentes tipos de erro possuem impactos assimétricos. A precisão quantifica a confiabilidade das predições positivas realizadas pelo modelo, sendo definida como a razão entre os verdadeiros positivos e o total de amostras classificadas como positivas:

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

Por sua vez, a sensibilidade — também denominada *recall* ou taxa de verdadeiros positivos — expressa a capacidade do modelo em identificar corretamente os indivíduos que pertencem à classe positiva, sendo calculada como:

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

A especificidade, ou taxa de verdadeiros negativos, quantifica a capacidade do modelo em identificar corretamente indivíduos pertencentes à classe negativa, sendo definida por:

$$Specificity = \frac{TN}{TN + FP} \quad (14)$$

Sob a ótica da Engenharia Elétrica, essas métricas podem ser interpretadas como parâmetros operacionais de um sistema de detecção sob incerteza, nos quais se estabelece um compromisso entre sensibilidade ao sinal de interesse e taxa de alarmes falsos.

#### 2.4.4 *F1-Score*

O *F1-score* constitui uma métrica composta amplamente empregada na avaliação de classificadores supervisionados, particularmente em cenários caracterizados por desbalanceamento entre classes e assimetria nos custos associados aos diferentes tipos de erro. Diferentemente da acurácia, que fornece uma medida global fortemente influenciada pela classe majoritária, o *F1-score* busca capturar, de forma equilibrada, a capacidade do modelo em identificar corretamente os casos positivos e a confiabilidade das predições realizadas para essa classe, conforme sistematizado por Powers [55].

Matematicamente, o *F1-score* é definido como a média harmônica entre a precisão (*Precision*) e a sensibilidade (*Recall*), sendo expresso por:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (15)$$

A escolha da média harmônica, em detrimento da média aritmética, não é arbitrária. Do ponto de vista matemático, a média harmônica penaliza fortemente valores extremos, assegurando que o *F1-score* assuma valores elevados apenas quando *Precision* e *Recall* apresentam simultaneamente bom desempenho. Assim, um classificador que obtenha alta sensibilidade à custa de um número excessivo de falsos positivos, ou elevada precisão acompanhada de baixa taxa de detecção de casos positivos, apresentará um *F1-score* reduzido. Essa propriedade torna a métrica particularmente adequada para problemas clínicos e de triagem, nos quais o equilíbrio entre detecção e confiabilidade é essencial.

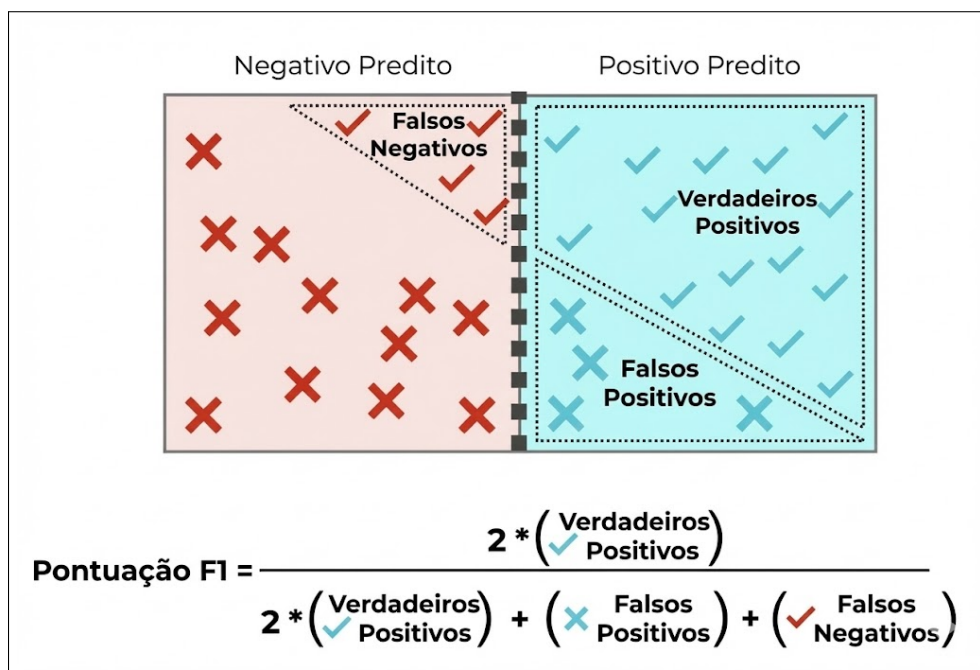
Sob a ótica da Engenharia Elétrica, o *F1-score* pode ser interpretado como uma medida escalar de desempenho global de um sistema de decisão binário que incorpora simultaneamente dois critérios potencialmente conflitantes: a capacidade de resposta do sistema (*recall*) e sua seletividade (*precision*). Essa interpretação é análoga à avaliação de sistemas de detecção operando em ambientes ruidosos, nos quais existe um compromisso inevitável entre sensibilidade à presença do sinal e taxa de alarmes falsos.

Em aplicações clínicas, como a triagem do risco de ansiedade pediátrica, esse compromisso assume relevância crítica. Uma sensibilidade elevada é desejável para minimizar a ocorrência de falsos negativos — isto é, crianças em situação de risco não identificadas pelo sistema —, enquanto uma precisão adequada é necessária para evitar a geração excessiva de alarmes falsos, que podem resultar em sobrecarga dos serviços de saúde e ansiedade desnecessária para

famílias e profissionais. O *F1-score*, ao sintetizar essas duas dimensões em uma única métrica, fornece uma avaliação mais justa e clinicamente informativa do desempenho do classificador.

Adicionalmente, o *F1-score* apresenta maior robustez frente ao desbalanceamento de classes quando comparado à acurácia, uma vez que ignora explicitamente os verdadeiros negativos em sua formulação. Essa característica é particularmente relevante em bases clínicas, nas quais a classe de maior interesse — como indivíduos com risco moderado ou elevado de ansiedade — tende a ser minoritária. Nesses cenários, o uso isolado da acurácia pode mascarar deficiências relevantes do modelo, enquanto o *F1-score* permanece sensível à qualidade da detecção dos casos clinicamente mais críticos.

**Figura 6 – Relação conceitual entre Precisão, Sensibilidade e *F1-Score***



Fonte: Representação conceitual adaptada pela autora a partir de [59].

A Figura 6 ilustra a interdependência entre precisão, sensibilidade e *F1-score*, destacando que essa métrica penaliza desequilíbrios entre seus componentes. Tal característica a torna especialmente adequada para contextos clínicos sensíveis, nos quais não é aceitável desempenho elevado em apenas um critério isolado. Neste estudo, o *F1-score* é utilizado como métrica central na avaliação dos classificadores, tanto em cenários binários quanto multiníveis, complementando a acurácia e proporcionando análise mais rigorosa sob condições de desbalanceamento.

## 2.4.5 Curva ROC e Área sob a Curva (AUC)

A curva *Receiver Operating Characteristic (ROC)*, formalizada e amplamente difundida na literatura de aprendizado de máquina por Fawcett [60], fornece uma representação gráfica do desempenho de um classificador ao longo de diferentes limiares de decisão, relacionando a taxa de verdadeiros positivos (*True Positive Rate – TPR*) e a taxa de falsos positivos (*False Positive Rate – FPR*). Essas taxas são definidas, respectivamente, como:

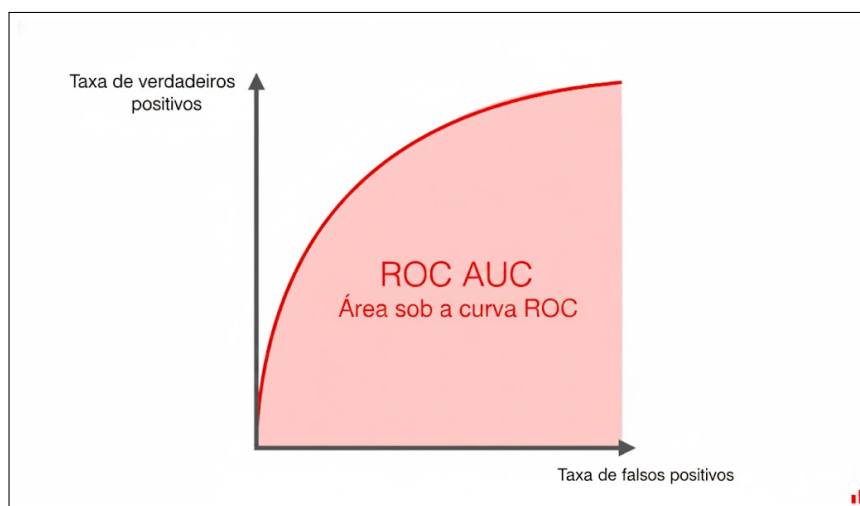
$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN} \quad (16)$$

A área sob a curva ROC, denominada *Area Under the Curve (AUC)*, sintetiza a capacidade discriminatória global do modelo, sendo interpretada como a probabilidade de o classificador atribuir uma pontuação maior a uma amostra positiva do que a uma negativa, conforme discutido por Fawcett [60]. Formalmente, a AUC pode ser expressa como:

$$AUC = \int_0^1 TPR(FPR) d(FPR) \quad (17)$$

Os valores de AUC situam-se no intervalo [0, 1], em que valores próximos de 1 indicam elevada separabilidade entre as classes, enquanto um valor de 0,5 corresponde a um desempenho equivalente ao de uma classificação aleatória.

**Figura 7 – Curva ROC e área sob a curva (AUC)**



Fonte: Adaptado pela autora a partir de [61].

A curva ROC explicita o compromisso (*trade-off*) entre sensibilidade e especificidade ao longo de diferentes limiares, enquanto a AUC fornece uma medida robusta e independente do limiar de decisão. Em aplicações clínicas, valores mais elevados de AUC tendem a indicar maior

separabilidade entre classes; entretanto, a interpretação do que constitui “bom desempenho” depende do contexto, do desbalanceamento entre classes e dos custos clínicos associados aos diferentes tipos de erro.

#### 2.4.6 Validação Cruzada Estratificada

A validação cruzada constitui um procedimento estatístico essencial para a estimação confiável do desempenho de modelos de aprendizado de máquina, particularmente em cenários caracterizados por bases de dados finitas, ruidosas e com distribuição de classes assimétrica. A fundamentação teórica desse procedimento é amplamente discutida por James et al. [62], bem como por Raschka [63]. Sob a perspectiva da Engenharia Elétrica, a validação cruzada pode ser interpretada como um método de estimação do erro médio de generalização de um sistema de decisão, análogo à avaliação do desempenho de sistemas dinâmicos sob múltiplas condições de excitação e diferentes realizações do sinal de entrada.

Entre as diversas variantes existentes, a validação cruzada estratificada (*Stratified Cross-Validation*) destaca-se por preservar, em cada partição do conjunto de dados, a proporção relativa das classes observada no conjunto original, conforme descrito na documentação técnica da biblioteca *scikit-learn* por Pedregosa et al. [64]. Essa propriedade é particularmente relevante em aplicações clínicas, nas quais o desbalanceamento entre classes é frequente e pode introduzir vieses significativos na estimação das métricas de desempenho, como sensibilidade, *F1-score* e erro clínico.

No método  $k$ -fold estratificado, o conjunto de dados  $\mathcal{D}$ , composto por  $N$  amostras, é particionado em  $k$  subconjuntos disjuntos  $\{\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_k\}$ , de forma que, a cada iteração, um subconjunto é utilizado como conjunto de validação, enquanto os  $k - 1$  subconjuntos restantes são empregados para o treinamento do modelo. A estratificação assegura que, para cada classe  $c \in \{1, \dots, C\}$ , a proporção de amostras seja aproximadamente mantida em todas as partições, conforme:

$$\frac{n_c^{(k)}}{n^{(k)}} \approx \frac{n_c}{N}, \quad \forall c \in \{1, \dots, C\}, \quad (18)$$

em que  $n_c$  representa o número total de amostras da classe  $c$  no conjunto completo,  $N$  o número total de amostras,  $n_c^{(k)}$  o número de amostras da classe  $c$  no  $k$ -ésimo *fold* e  $n^{(k)}$  o número total de amostras nesse subconjunto. Essa condição reduz a variância da estimativa do desempenho e evita que partições específicas sejam dominadas por classes majoritárias.

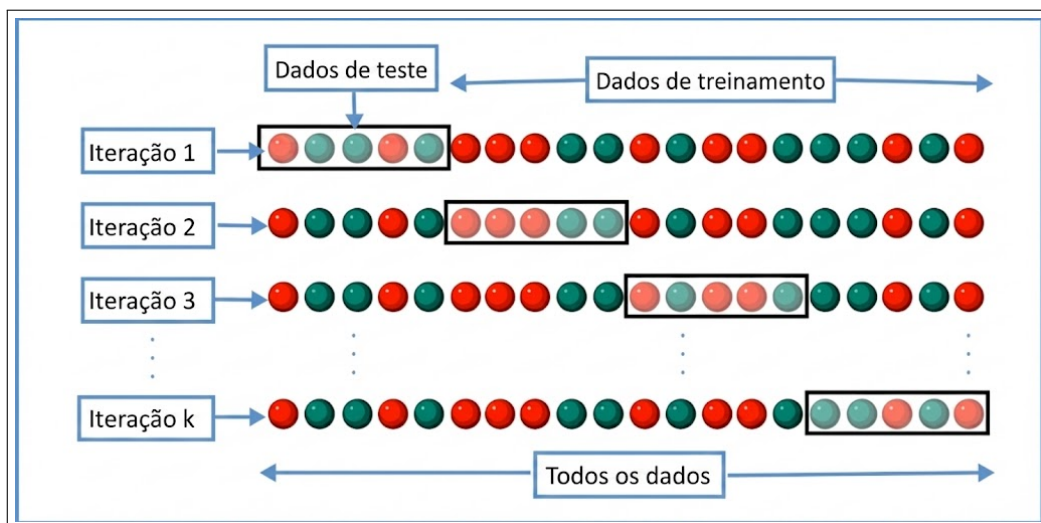
Do ponto de vista estatístico, a validação cruzada estratificada permite estimar o desempenho médio esperado do classificador por meio da média das métricas obtidas em cada iteração, formalmente expressa como:

$$\hat{\mathcal{M}} = \frac{1}{k} \sum_{i=1}^k \mathcal{M}_i, \quad (19)$$

em que  $\mathcal{M}_i$  representa a métrica de desempenho avaliada no  $i$ -ésimo *fold*. Essa estimativa apresenta menor viés em relação a uma única partição treino-teste, sendo particularmente indicada para bases clínicas de tamanho moderado, conforme discutido por James et al. [62] e Max Kuhn e Julia Silge [65].

Sob a ótica da Engenharia Elétrica aplicada à saúde, a validação cruzada estratificada pode ser interpretada como um mecanismo de teste de robustez do sistema de decisão frente a variações estatísticas do conjunto de entrada. Ao garantir que todas as classes de risco estejam representadas de forma consistente em cada iteração, o método contribui para avaliações mais estáveis, reproduzíveis e clinicamente confiáveis.

**Figura 8 – Representação conceitual da validação cruzada estratificada**



Fonte: Adaptado pela autora a partir de [66].

No presente trabalho, a validação cruzada estratificada é empregada tanto na etapa de otimização de hiperparâmetros quanto na avaliação final dos modelos *Random Forest*, *Support Vector Machines* e *Multilayer Perceptron*. Essa escolha metodológica assegura consistência estatística, reduz o risco de viés associado ao desbalanceamento de classes e fortalece a validade clínica das inferências produzidas pelo sistema proposto.

## Síntese Final do Referencial Teórico

À luz das discussões apresentadas, a ansiedade pediátrica pode ser formalizada como um problema de inferência de estados latentes contínuos a partir de observações discretas, ruidosas, subjetivas e parcialmente observáveis, caracterizadas por elevada variabilidade interindividual e organização em espaços de alta dimensionalidade, conforme discutido por Kendall [32] e Costello et al. [33]. A natureza multifatorial, não linear e ordinal do risco clínico — em consonância com a abordagem dimensional proposta por Kraemer [39] — impõe limitações a abordagens analíticas simplificadas, demandando modelos capazes de operar sob incerteza estatística e estrutural, e capturar dependências complexas entre múltiplas variáveis.

Sob a perspectiva da Engenharia Elétrica, esse enquadramento aproxima o problema de paradigmas clássicos de processamento de sinais e reconhecimento de padrões, conforme sistematizado por Bishop [7] e por Hastie [8], nos quais o objetivo consiste em extrair representações informativas a partir de dados imperfeitos e inferir estados subjacentes não diretamente observáveis. Nesse contexto, modelos computacionais baseados em aprendizado supervisionado constituem ferramentas teóricas adequadas para a modelagem de fenômenos psicométricos complexos, ao possibilitar a aprendizagem de relações não lineares e a estimação de fronteiras de decisão compatíveis com a natureza ordinal do risco clínico.

O avanço do aprendizado de máquina tem impulsionado a psiquiatria computacional, integrando modelos matemáticos ao raciocínio clínico, conforme discutido por Shatte et al. [6]. Arquiteturas consolidadas — como o *Random Forest* proposto por Breiman, as *Support Vector Machines* de Cortes e Vapnik e as redes neurais artificiais sistematizadas por Goodfellow et al. — destacam-se pela capacidade de modelar relações não lineares em dados psicométricos tabulares aplicados a contextos clínicos.

Adicionalmente, os referenciais diagnósticos normativos, DSM-5 [10] e a CID-11 [11], fornecem o enquadramento clínico necessário para a interpretação dos construtos avaliados, embora não eliminem a incerteza inerente aos instrumentos de avaliação e à atribuição de rótulos clínicos. Nesse cenário, a incorporação de métodos de explicabilidade, como o *SHAP*, proposto por Lundberg e Lee [54], torna-se fundamental para assegurar transparência, auditabilidade e coerência clínica das decisões dos modelos. O referencial teórico sustenta as escolhas metodológicas e a modelagem do risco de ansiedade pediátrica. A abordagem mantém alinhamento com a Engenharia Elétrica aplicada à saúde mental, priorizando robustez estatística, interpretabilidade e segurança clínica, conforme detalhado no capítulo seguinte.

### 3 METODOLOGIA

Esta seção descreve, de forma sistemática e detalhada, os fundamentos conceituais, os procedimentos computacionais e as técnicas analíticas adotadas no desenvolvimento do presente estudo. A metodologia foi concebida com ênfase em rigor estatístico, reprodutibilidade experimental e coerência epistemológica, atendendo às exigências metodológicas de pesquisas desenvolvidas no âmbito da Engenharia Elétrica aplicada à saúde.

O delineamento metodológico foi estruturado de modo a permitir a avaliação comparativa de diferentes arquiteturas de aprendizado supervisionado, considerando explicitamente seus perfis decisórios em termos de sensibilidade, estabilidade de generalização e impacto clínico do erro. Foram investigados os modelos *Random Forest (RF)*, *Support Vector Machine (SVM)* e redes neurais artificiais do tipo *Multilayer Perceptron (MLP)*, avaliando sua capacidade de estimar o risco de transtornos de ansiedade pediátrica em cenários binário e multiclasse ordinal. Esses modelos foram selecionados por representarem abordagens complementares de modelagem estatística e conexionista, amplamente consolidadas na literatura de sistemas inteligentes, processamento estatístico de sinais e aprendizado de máquina aplicado a problemas complexos e ruidosos.

Sob a perspectiva da Engenharia Elétrica, o problema é formulado como um sistema de decisão supervisionado operando sobre vetores de atributos psicométricos discretos, nos quais cada vetor de entrada representa observações ruidosas de estados emocionais latentes. Nesse enquadramento, cada amostra do conjunto de dados é interpretada como um vetor de observações associado a manifestações comportamentais e emocionais, enquanto os modelos supervisionados atuam como sistemas decisórios não lineares responsáveis por inferir estados latentes de risco clínico. Tal formulação permite interpretar o processo de aprendizado como uma etapa de extração de padrões relevantes em presença de ruído, variabilidade interindividual e sobreposição estatística entre classes.

Cabe ressaltar que, neste trabalho, o termo “detecção” é empregado em sentido abstrato, referindo-se à identificação de padrões de risco no espaço de atributos psicométricos, e não à detecção clássica de sinais fisiológicos contínuos no domínio temporal. A escolha dessa abordagem metodológica justifica-se pela natureza intrinsecamente multivariada, não linear e contínua dos fenômenos emocionais na infância. A expressão clínica da ansiedade pediátrica envolve interações complexas entre dimensões afetivas, cognitivas e comportamentais, cujas fronteiras entre níveis de severidade são difusas e frequentemente sobrepostas. A modelagem adequada

desse tipo de fenômeno demanda técnicas capazes de representar dependências latentes e relações não lineares, características centrais dos modelos supervisionados adotados neste trabalho.

Todo o estudo foi desenvolvido integralmente em ambiente *in silico*, isto é, por meio de experimentação computacional, sem coleta primária de dados clínicos ou interação direta com participantes humanos. Essa abordagem assegura rigor ético, controle experimental e elevada reprodutibilidade dos resultados. As etapas do pipeline metodológico — incluindo pré-processamento dos dados, definição dos cenários experimentais, treinamento dos modelos, validação estatística, avaliação de desempenho e análise interpretativa — foram integralmente implementadas em linguagem Python, com o uso de bibliotecas científicas consolidadas para análise de dados, aprendizado de máquina e visualização.

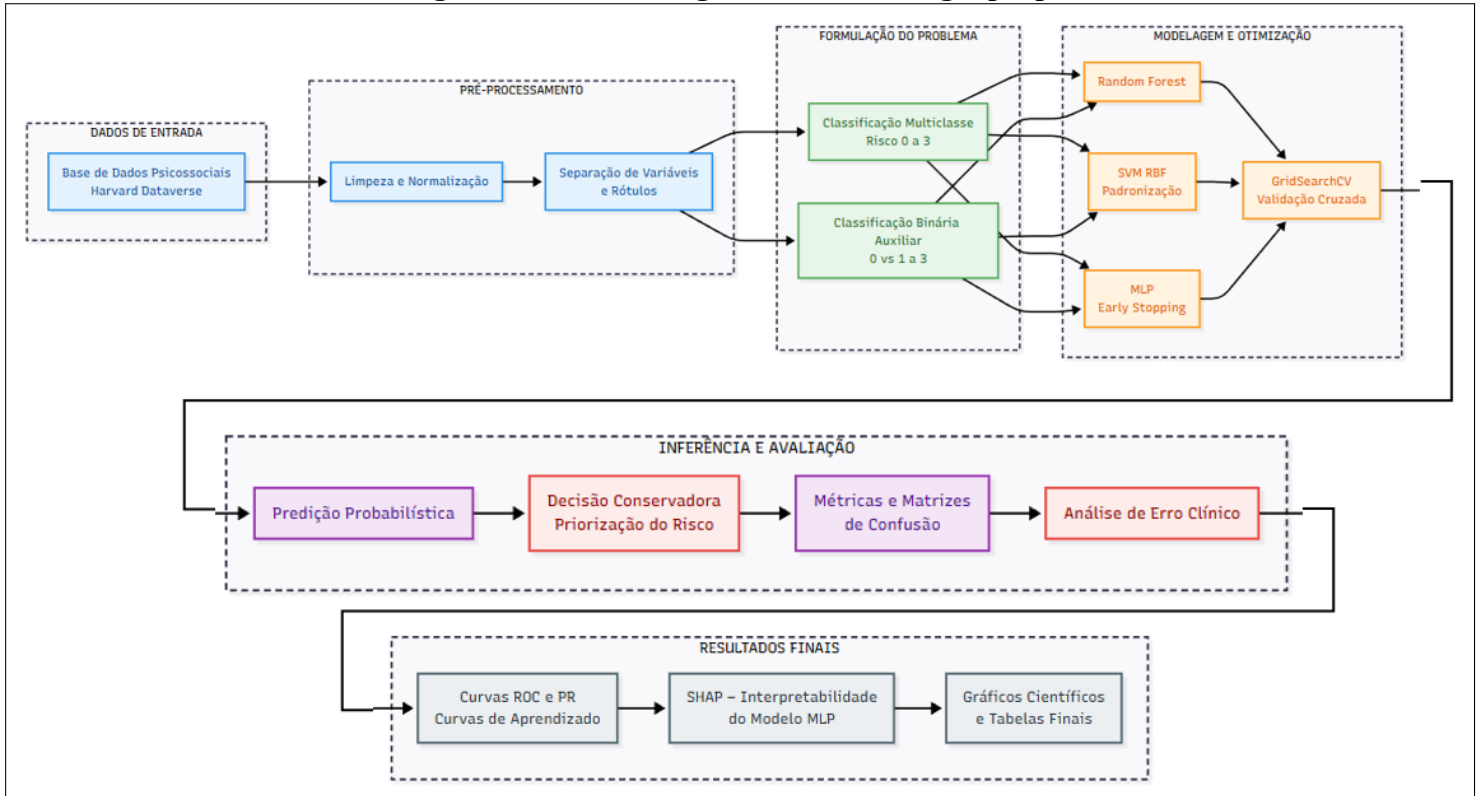
O fluxo geral da metodologia adotada é apresentado na Figura 9, a qual sintetiza as principais etapas do processo experimental. As subseções seguintes detalham cada componente desse pipeline, descrevendo a base de dados utilizada, os procedimentos de pré-processamento, os critérios de configuração dos modelos, as estratégias de validação e os métodos empregados para análise quantitativa e clínica dos resultados.

### 3.1 Estrutura Geral da Metodologia

A metodologia adotada neste trabalho é de natureza quantitativa, experimental e computacional, sendo estruturada segundo um pipeline clássico de sistemas inteligentes aplicados à classificação de risco em saúde. O delineamento metodológico foi concebido para modelar, de forma sistemática, o processo de inferência do risco de ansiedade pediátrica a partir de dados psicométricos e comportamentais, integrando etapas de pré-processamento, modelagem supervisionada, inferência probabilística e avaliação clínica do erro.

Sob a perspectiva da Engenharia Elétrica, o problema é formulado como um sistema de decisão supervisionado operando sobre vetores psicométricos discretos, nos quais cada vetor de entrada representa observações ruidosas de estados emocionais latentes. O pipeline metodológico foi organizado de modo a permitir a avaliação comparativa de diferentes arquiteturas de aprendizado supervisionado — *Random Forest (RF)*, *Support Vector Machine (SVM)* e *Multilayer Perceptron (MLP)* — considerando não apenas métricas estatísticas tradicionais, mas também critérios de impacto clínico e segurança decisória. A Figura 9 apresenta a estrutura geral da metodologia proposta, evidenciando as principais etapas do processo experimental e suas inter-relações. O pipeline é composto por cinco macroblocos funcionais, descritos a seguir.

**Figura 9 – Estrutura geral da metodologia proposta**



Fonte: Elaboração própria (2026).

Esse encadeamento metodológico assegura consistência entre formulação do problema, implementação computacional e critérios de avaliação, servindo como base para as análises apresentadas nas seções subsequentes.

## 3.2 Materiais

Esta subseção descreve os materiais empíricos e computacionais empregados no desenvolvimento do estudo, contemplando a origem dos dados, suas características estruturais e clínicas, bem como os recursos computacionais utilizados para a execução dos experimentos. A seleção desses materiais foi orientada por critérios de validade científica, reprodutibilidade experimental e aderência ética, fundamentais para pesquisas em saúde mental pediátrica sob a perspectiva da Engenharia Elétrica.

### 3.2.1 Fonte de Dados: *Harvard Dataverse*

A base empírica utilizada neste estudo foi obtida a partir do repositório público *Harvard Dataverse*, originalmente disponibilizado por Carpenter et al. [67], reunindo informações psicométricas, comportamentais e demográficas provenientes de estudos conduzidos com crianças

em idade pré-escolar (3 a 5 anos).

Os dados foram coletados no contexto do estudo da autora supracitada, cujo objetivo consistiu em investigar o uso de técnicas de aprendizado de máquina para estimativa de risco de transtornos de ansiedade na infância, com ênfase na redução do número de itens clínicos necessários para predição diagnóstica. No estudo original, foi empregada exclusivamente a técnica simbólica *Alternating Decision Tree (ADTree)*, sem comparação sistemática com outros paradigmas supervisionados (por exemplo, *ensembles*, modelos de margem ou redes neurais).

No presente trabalho, o *dataset* é utilizado como base para avaliar, de forma comparativa, modelos supervisionados com perfis decisórios complementares — como *Random Forest*, *Support Vector Machines (SVM)* e *Multilayer Perceptron (MLP)* — no contexto de classificação binária e multiclasse ordinal do risco de ansiedade pediátrica, ampliando metodologicamente o escopo originalmente proposto por Carpenter et al.

### **I - Estrutura dos Arquivos**

O conjunto de dados disponibilizado no *Harvard Dataverse* é composto por dois arquivos principais, organizados conforme a divisão original proposta pelos autores. Na versão utilizada neste estudo, o arquivo *Training Data.xlsx* contém 918 registros independentes, enquanto o arquivo *Testing Data.xlsx* reúne 407 registros, totalizando 1325 crianças avaliadas. Embora o estudo original disponibilize essa separação pré-definida, neste trabalho optou-se por consolidar ambos os arquivos. Assim, a separação original (*Training/Testing*) foi utilizada apenas como organização de disponibilização, não como protocolo experimental no presente estudo.

Após as etapas de organização, verificação estrutural e tratamento dos dados, obteve-se uma base consolidada composta por 872 registros válidos para modelagem supervisionada. A redução do número inicial de registros decorreu da remoção de observações com inconsistências estruturais entre atributos, dados incompletos em variáveis críticas e eventuais problemas identificados na etapa de verificação de integridade. Os critérios de exclusão foram aplicados de forma determinística e previamente definidos, priorizando a preservação de consistência entre atributos, completude das variáveis críticas e integridade da variável-alvo.

Cada registro corresponde a uma criança distinta, não havendo múltiplas observações repetidas de um mesmo indivíduo no nível de modelagem adotado neste estudo. A partir dessa base consolidada, foi aplicado particionamento estratificado próprio (70%/30%), resultando em 610 instâncias no conjunto de treinamento e 262 instâncias no conjunto de teste.

## II - Tipos de Variáveis

A base apresenta estrutura multivariada e heterogênea, integrando diferentes domínios relevantes para a modelagem do risco de ansiedade pediátrica. Entre as variáveis demográficas, incluem-se idade, sexo, etnia e indicadores socioeconômicos, incluindo o nível socioeconômico (*Socioeconomic Status – SES*).

As medidas psicométricas provêm de instrumentos clinicamente validados, com destaque para *BAS-B* e *BAS-F*, relacionadas à inibição comportamental e respostas de medo; *ARI*, associada à inibição relacionada à ansiedade; *ERC-LER* e *ERC-ERS*, referentes à labilidade e à regulação emocional; e *SDQ-ES*, subescala de sintomas emocionais do *Strengths and Difficulties Questionnaire (SDQ)*.

A base inclui ainda medidas psicofisiológicas, como frequência cardíaca média, condutância da pele (*Skin Conductance Level – SCL*) e reatividade do cortisol, fornecendo indicadores indiretos de ativação autonômica e estresse fisiológico. Também estão disponíveis indicadores diagnósticos binários associados à presença/ausência de condições psiquiátricas, incluindo GAD, SAD, fobia social, depressão, TDAH, TOD e TC, bem como probabilidades contínuas baseadas em distribuição Gamma, correspondentes a estimativas probabilísticas de risco para GAD e SAD, modeladas por distribuição Gamma e representativas de assimetria e variabilidade do risco clínico. Adicionalmente, há um diagnóstico longitudinal, obtido em acompanhamento após 12 meses, utilizada como referência clínica adicional.

Embora a base inclua variáveis psicofisiológicas, essas medidas são tratadas neste trabalho como atributos agregados em formato tabular, não sendo analisadas como sinais fisiológicos contínuos no domínio temporal.

## III - Características Gerais da Base

O dataset apresenta propriedades particularmente adequadas ao escopo deste estudo. Destaca-se que a amostra é composta exclusivamente por crianças em idade pré-escolar, alinhada ao objetivo de triagem precoce. Além disso, a base possui estrutura multimodal (psicométrica, comportamental e psicofisiológica) em formato tabular, apresenta baixa incidência de valores ausentes — o que reduz a necessidade de imputação complexa — e oferece estrutura adequada para modelagem supervisionada e validação estatística controlada. Por fim, trata-se de um conjunto de dados de disponibilidade pública sob licença CC0 (*Creative Commons Zero*), garantindo uso irrestrito, anonimização e conformidade ética.

## IV - Relevância da Base para o Estudo

A adoção dessa base fundamenta-se em quatro aspectos. Primeiro, sua validade clínica, em razão do uso de instrumentos padronizados e avaliações conduzidas sob protocolos reconhecidos. Segundo, sua aderência ao problema investigado, uma vez que o dataset foi concebido especificamente para a modelagem do risco de ansiedade pediátrica. Terceiro, sua compatibilidade com modelos não lineares, apresentando estrutura adequada para *RF*, *SVM* e *MLP*. Por fim, sua conformidade ética e científica, por se tratar de dados públicos, anonimizados e reprodutíveis. Dessa forma, o conjunto de dados do *Harvard Dataverse* [67] constitui um recurso empiricamente robusto e clinicamente relevante, fornecendo base consistente para o desenvolvimento, treinamento e avaliação dos modelos supervisionados propostos neste trabalho.

### 3.2.2 Ambiente Computacional e Ferramentas de Software

Os experimentos foram conduzidos em Python 3.10, em sistema operacional Windows 11, em computador com processador Intel i7 e 16 GB de memória RAM. O desenvolvimento foi realizado no *Visual Studio Code (VS Code)*, com uso de notebooks e extensões para Python, favorecendo documentação contínua, rastreabilidade do código e reprodutibilidade experimental.

As bibliotecas empregadas incluem *TensorFlow/Keras*, para implementação e treinamento do *Multilayer Perceptron (MLP)*, incluindo regularização e *early stopping*; *Scikit-learn*, para pré-processamento, implementação de *RF* e *SVM*, validação cruzada estratificada, otimização via *GridSearchCV* e cálculo de métricas; *Pandas* e *NumPy*, para manipulação tabular, processamento vetorial e operações numéricas; *Matplotlib*, para geração de gráficos científicos (curvas ROC, *Precision–Recall* e curvas de aprendizado); *SHAP*, para análise de interpretabilidade do modelo *MLP* por valores de Shapley; e *OpenPyXL*, para leitura e escrita de arquivos *.xlsx*. A adoção de bibliotecas de código aberto reforça a transparência metodológica e a reprodutibilidade científica, em consonância com os princípios de *Open Science* propostos por Wilkinson et al. [68].

### 3.2.3 Organização e Tratamento dos Dados

Os dados foram submetidos a um processo sistemático de curadoria e pré-processamento com o objetivo de assegurar consistência estatística, compatibilidade entre os conjuntos de treinamento e teste e adequação às arquiteturas supervisionadas investigadas. Todas as etapas foram

conduzidas de forma determinística, garantindo rastreabilidade e controle experimental durante o processo de consolidação e transformação dos dados.

O processo de organização e tratamento dos dados compreendeu as seguintes etapas:

### **I – Padronização estrutural e conversão numérica**

Inicialmente, realizou-se a verificação da consistência estrutural dos atributos, bem como a conversão de todas as variáveis para formato numérico compatível com os algoritmos de aprendizado de máquina empregados. Essa etapa garantiu a integridade dos vetores de entrada e a uniformidade da representação dos dados.

### **II – Tratamento de valores ausentes**

Em razão da baixa incidência de valores faltantes na base, optou-se por uma estratégia de imputação simples e determinística. Essa escolha assegurou compatibilidade com os modelos supervisionados adotados, sem introduzir viés estatístico significativo ou complexidade adicional ao processo de modelagem.

### **III – Normalização dos atributos**

Para mitigar efeitos de escala entre variáveis de diferentes naturezas, foi aplicada a padronização dos atributos contínuos por meio do método `StandardScaler()`, garantindo média zero e variância unitária. Essa etapa é particularmente relevante em modelos sensíveis à magnitude dos atributos, como *SVM* e *MLP*, além de favorecer a estabilidade do processo de treinamento.

### **IV – Definição das variáveis-alvo**

Por fim, foi construída a variável-alvo `Anxiety_Multilevel`, categorizada em quatro níveis ordinais de risco (0, 1, 2 e 3), representando o cenário multiclasse do problema. Adicionalmente, foi derivado um cenário binário auxiliar (0 vs. 1–3), utilizado para análises complementares de desempenho e comparação com abordagens de triagem clínica. Ao término dessas etapas, obteve-se uma base de dados limpa, normalizada e estruturalmente consistente, consolidada em arquivo intermediário e utilizada como entrada principal nos experimentos de modelagem supervisionada. Esse procedimento garante coerência entre as fases de treinamento e avaliação, além de sustentar a validade metodológica dos resultados apresentados nas seções subsequentes.

### **3.3 Métodos**

Esta subseção descreve os métodos computacionais e estatísticos empregados no desenvolvimento do sistema de apoio à decisão proposto. O procedimento foi organizado em etapas sequenciais, de modo a refletir fielmente o pipeline implementado (Figura 9) e assegurar reprodutibilidade, rigor estatístico e coerência clínica na avaliação dos modelos. O processo metodológico foi estruturado nas etapas descritas a seguir.

#### **3.3.1 Etapa 1 – Formulação do Problema e Definição dos Cenários Experimentais**

O risco de ansiedade pediátrica foi modelado como um problema de classificação supervisionada multiclasse ordinal, no qual a variável-alvo assume quatro níveis discretos e ordenados de risco clínico,  $\{0, 1, 2, 3\}$ . Essa formulação é consistente com a natureza dimensional do fenômeno e permite analisar não apenas acertos, mas também a gravidade semântica dos erros entre níveis de risco.

Adicionalmente, foi definido um cenário binário auxiliar (0 vs. 1–3), no qual a classe 0 representa ausência de risco e as classes 1, 2 e 3 são agrupadas como presença de risco. Esse cenário foi empregado para análises complementares de triagem, permitindo o uso de métricas clássicas frequentemente utilizadas em aplicações clínicas, como sensibilidade e especificidade.

#### **3.3.2 Etapa 2 – Pré-processamento e Preparação dos Dados**

Os dados passaram por curadoria e pré-processamento com o objetivo de garantir consistência estatística e compatibilidade com os algoritmos supervisionados. As variáveis explicativas foram convertidas para formato numérico e os valores ausentes foram tratados por imputação simples pela média, com parâmetros estimados exclusivamente no conjunto de treinamento e posteriormente aplicados ao conjunto de teste, de forma análoga ao procedimento de controle de vazamento de informação descrito na Etapa 3.

A etapa de normalização por `StandardScaler()` foi planejada nesta fase e executada de forma a evitar vazamento de informação, com ajuste dos parâmetros exclusivamente no conjunto de treinamento e aplicação posterior ao conjunto de teste, conforme detalhado na Etapa 3.

#### **3.3.3 Etapa 3 – Particionamento e Controle de Vazamento de Informação**

O particionamento dos dados seguiu uma divisão treino–teste de 70%/30%, com amostragem estratificada em relação à variável-alvo, preservando a distribuição das classes em ambas as

partições. O particionamento foi aplicado à base consolidada após as etapas de curadoria estrutural e definição da variável-alvo. As etapas de padronização (*StandardScaler*) e demais transformações dependentes de parâmetros estatísticos foram ajustadas exclusivamente no conjunto de treinamento e posteriormente aplicadas ao conjunto de teste, de modo a evitar vazamento de informação (*data leakage*), resultando em  $N_{\text{treino}} = 610$  e  $N_{\text{teste}} = 262$  instâncias. O conjunto de teste permaneceu isolado durante todo o processo de ajuste e seleção de hiperparâmetros, sendo utilizado exclusivamente para avaliação final, enquanto a validação cruzada estratificada foi empregada apenas no conjunto de treinamento, reduzindo o risco de estimativas otimistas de desempenho.

### 3.3.4 Etapa 4 – Seleção e Configuração dos Modelos Supervisionados

Foram avaliados três classificadores supervisionados, selecionados por sua consolidação na literatura e por representarem paradigmas complementares de decisão:

- **Random Forest (RF)**: classificador por comitê de árvores, robusto a ruído e adequado a relações não lineares locais;
- **Support Vector Machine (SVM)** com kernel *RBF*: modelo de margem com fronteiras não lineares, implementado em *pipeline* com padronização explícita;
- **Multilayer Perceptron (MLP)**: rede neural totalmente conectada, com alta capacidade representacional para modelar interações não lineares entre atributos.

Sob a ótica da Engenharia Elétrica, esses modelos podem ser interpretados como sistemas decisórios não lineares operando sobre vetores de atributos psicométricos discretos, inferindo estados latentes de risco clínico a partir de observações ruidosas e parcialmente observáveis.

### 3.3.5 Etapa 5 – Otimização de Hiperparâmetros por Validação Cruzada

A otimização dos hiperparâmetros foi realizada por *GridSearchCV* com validação cruzada estratificada em cinco dobras (*StratifiedKFold*,  $k = 5$ ), permitindo estimar desempenho médio de forma mais robusta e reduzindo dependência de uma partição específica. Foram explorados, de forma sistemática, parâmetros típicos de cada modelo (por exemplo, número de

árvores e profundidade no *RF*;  $C$  e  $\gamma$  no *SVM*; arquitetura, função de ativação, taxa de aprendizado e regularização no *MLP*. O critério de seleção adotado foi a maximização da acurácia média em validação cruzada, garantindo comparabilidade entre os modelos no cenário multiclasse.

A escolha da acurácia como critério primário de otimização decorreu da necessidade de comparabilidade direta entre os modelos no cenário multiclasse, sendo as demais métricas (F1-score, sensibilidade e especificidade) utilizadas posteriormente para avaliação complementar do comportamento decisório.

### 3.3.6 Etapa 6 – Inferência Probabilística e Estratégia de Decisão Conservadora

Após o ajuste final, os modelos foram aplicados ao conjunto de teste com geração de probabilidades por classe, permitindo avaliar o comportamento decisório sob incerteza classificatória. Com base nessas probabilidades, foi adotada uma estratégia de decisão conservadora no cenário multiclasse ordinal. Em situações de baixa confiança preditiva, caracterizadas por baixa separação entre as probabilidades das duas classes mais prováveis, a decisão foi direcionada para a classe de maior risco entre as candidatas. Operacionalmente, a baixa confiança foi definida por um critério baseado na diferença  $\Delta p = p_{(1)} - p_{(2)}$  entre as duas maiores probabilidades estimadas, onde  $p_{(1)}$  representa a maior probabilidade predita e  $p_{(2)}$  a segunda maior. Esse critério foi aplicado de forma determinística e reproduzível.

O limiar de diferença probabilística adotado foi estimado exclusivamente com base na distribuição das probabilidades observadas no conjunto de treinamento, utilizando critério estatístico previamente definido. O valor obtido foi mantido fixo para todos os modelos e aplicado de forma idêntica ao conjunto de teste, sem qualquer ajuste posterior. Tal procedimento busca reduzir a probabilidade de subestimação do nível de risco em condições de ambiguidade preditiva. A estratégia foi concebida por analogia a sistemas decisórios críticos, nos quais a assimetria de custo associada à subestimação pode ser mais relevante do que a superestimação, especialmente em contextos de triagem clínica.

### 3.3.7 Etapa 7 – Avaliação de Desempenho e Análise de Erro Clínico Ordinal

A avaliação final foi conduzida no conjunto de teste, contemplando os cenários multiclasse e binário.

No cenário multiclasse (0–3), foram calculadas métricas globais e *macro-averaged* (acurácia, precisão, sensibilidade, especificidade e *F1-score*), além de *AUC-ROC* por estratégia *one-*

*vs-rest* e matriz de confusão multiclasse. O uso de medidas *macro* foi adotado para mitigar efeitos do desbalanceamento entre níveis de risco.

Além das métricas estatísticas, foi realizada a análise de erro clínico ordinal, categorizando erros em:

- **Erros adjacentes:** confusão entre classes vizinhas, com menor impacto clínico esperado;
- **Erros não adjacentes:** discrepância de dois ou mais níveis, associada a subestimação ou superestimação clinicamente relevante.

No cenário binário (0 vs. 1–3), foram computadas acurácia, sensibilidade, especificidade, precisão, *F1-score* e *AUC-ROC*, permitindo leitura complementar do desempenho em contexto de triagem.

### 3.3.8 Etapa 8 – Interpretabilidade do Modelo por *SHAP*

A análise de interpretabilidade foi conduzida prioritariamente sobre o modelo *MLP*, em razão de sua natureza conexionista e maior complexidade representacional em comparação aos modelos baseados em árvores e margens. Para tal, foram calculados valores *SHAP* (*SHapley Additive Explanations*), os quais quantificam a contribuição marginal de cada variável de entrada para a predição do modelo em cada instância avaliada.

A interpretabilidade foi reportada por meio de métricas globais de importância (impacto médio absoluto dos valores de Shapley), permitindo identificar os atributos mais influentes na estratificação do risco. Essa análise fornece subsídios para transparência do processo decisório e para discussão interdisciplinar dos resultados no contexto clínico.

### 3.3.9 Síntese Metodológica

Em síntese, a metodologia foi estruturada como um pipeline integrado que combina pré-processamento padronizado, comparação de modelos supervisionados, otimização por validação cruzada, inferência probabilística orientada ao risco, avaliação estatística e análise de erro clínico ordinal, complementadas por interpretabilidade via *SHAP*. Esse encadeamento fornece base metodológica robusta para a análise comparativa apresentada nos capítulos subsequentes e para a discussão do classificador mais adequado como núcleo do sistema de apoio à decisão proposto.

## 4 RESULTADOS

Esta seção apresenta os resultados obtidos a partir da implementação e avaliação dos modelos computacionais desenvolvidos neste trabalho, aplicados à base de dados do *Harvard Dataverse*, descrita por Carpenter et al. [9]. Os arquivos originais (*Training/Testing*) foram consolidados e, após as etapas de curadoria, verificação de integridade e pré-processamento, obteve-se uma base final composta por  $N = 872$  registros válidos para modelagem. Cada registro corresponde a uma criança distinta, não havendo repetição de indivíduos no conjunto consolidado final. A partir dessa base, foi aplicado particionamento estratificado próprio (70%/30%), resultando em  $N_{\text{treino}} = 610$  e  $N_{\text{teste}} = 262$ . Os resultados foram organizados em dois cenários principais: **Cenário Binário (0 vs 1–3)** – distinção entre ausência de risco e presença de risco; e **Cenário Multiclasse (0–3)** – classificação detalhada em quatro níveis ordenados de risco.

### 4.1 Desempenho Binário (0 vs 1–3)

Neste cenário, o problema de classificação foi reformulado para distinguir registros (crianças) sem risco (classe 0) daqueles com risco (classes 1, 2 e 3 agregadas). Essa configuração é típica de sistemas de triagem clínica, cujo objetivo é identificar casos que demandam encaminhamento para avaliação especializada, independentemente do grau de severidade.

#### 4.1.1 Matriz de Confusão

A Tabela 1 apresenta os valores extraídos das matrizes de confusão binárias, detalhando os acertos e erros de cada modelo.

Tabela 1: **Distribuição dos acertos e erros no cenário binário (0 vs 1–3).**

Modelo	VN	FP	FN	VP	Total ( $N_{\text{teste}}$ )	Acerto (%)
<i>SVM</i>	142	12	16	92	262	89,31
<i>MLP</i>	139	15	20	88	262	86,64
<i>Random Forest</i>	149	5	20	88	262	90,46

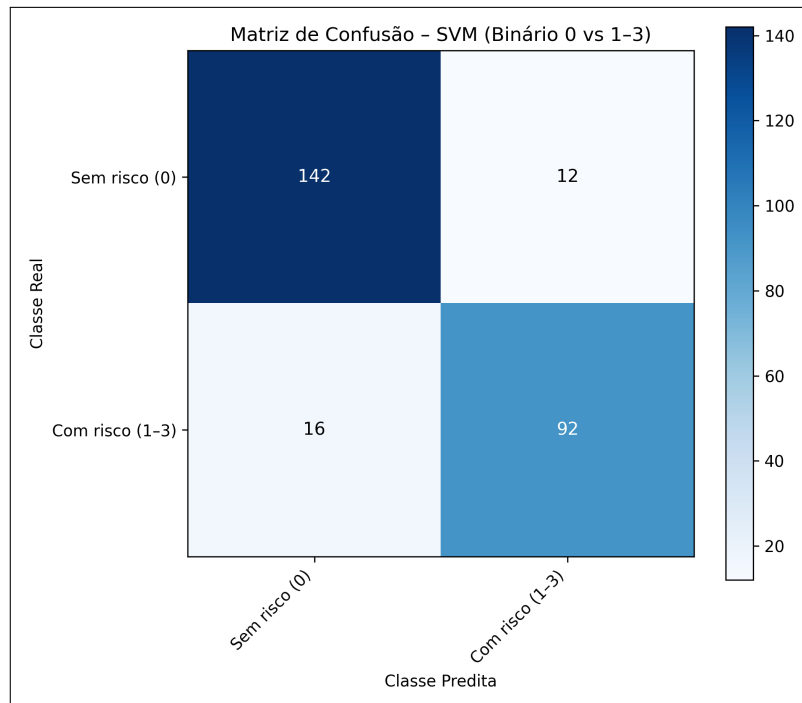
**Nota:** VN = verdadeiros negativos; FP = falsos positivos; FN = falsos negativos; VP = verdadeiros positivos. Os valores referem-se ao desempenho no conjunto de teste ( $N_{\text{teste}} = 262$  registros).

**Fonte:** Elaboração própria (2026).

O *RF* apresenta perfil mais conservador, com maior número de verdadeiros negativos e menor taxa de falsos positivos (maior especificidade). O *SVM* mostra equilíbrio entre sensibili-

dade e especificidade, enquanto o *MLP* mantém desempenho consistente, com elevado número de verdadeiros positivos.

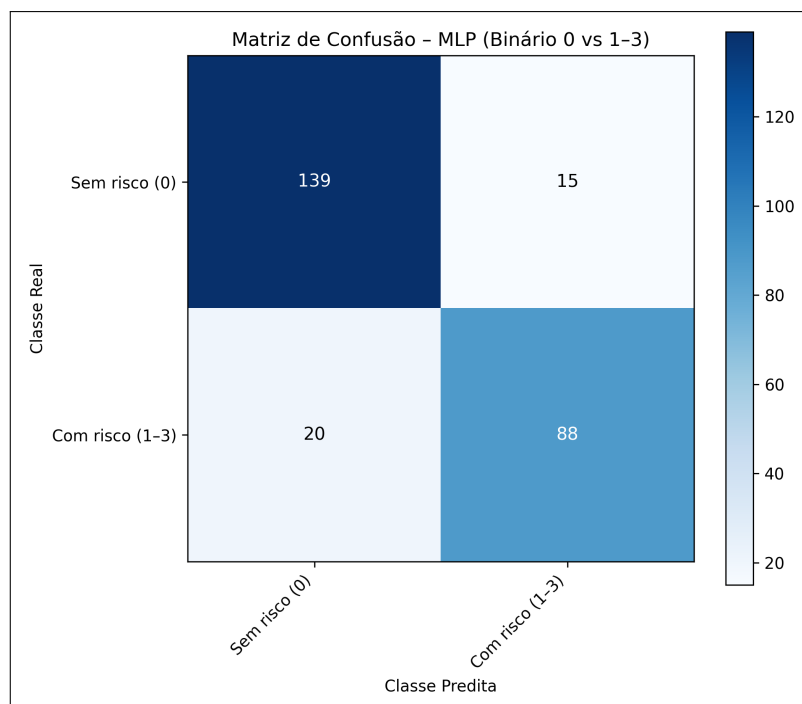
**Figura 10 – Matriz de Confusão Binária do modelo *SVM* (0 vs. 1-3)**



Fonte: Elaboração própria (2026).

Em particular, o *SVM* apresenta um comportamento mais equilibrado entre erros do tipo falso positivo e falso negativo, sugerindo compromisso estável entre sensibilidade e especificidade.

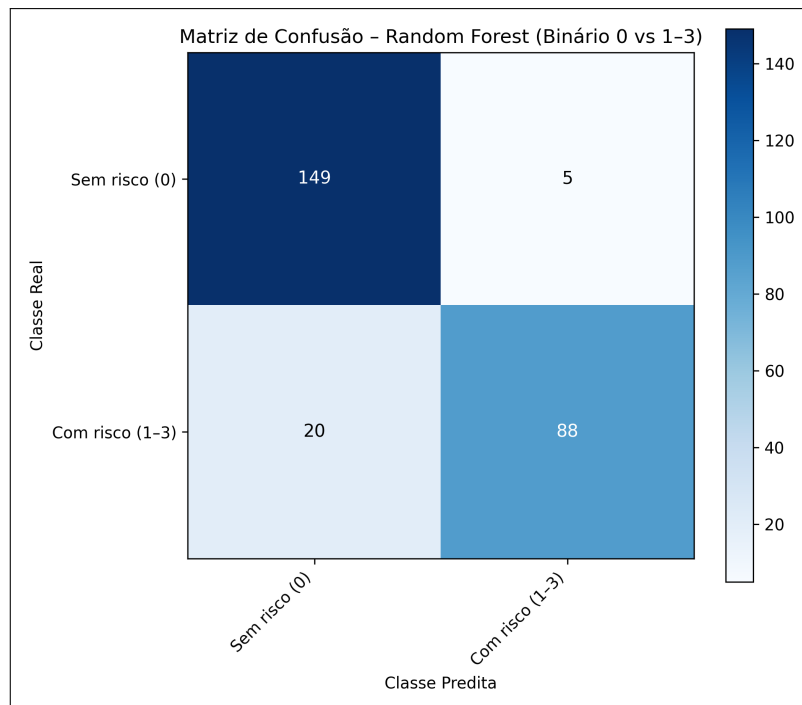
**Figura 11 – Matriz de Confusão Binária do modelo *MLP* (0 vs. 1-3)**



Fonte: Elaboração própria (2026).

O *MLP* mantém desempenho consistente, com elevada capacidade de detecção de casos de risco, refletindo sua flexibilidade para modelar relações não lineares em dados tabulares psicométricos.

**Figura 12 – Matriz de Confusão Binária do modelo *Random Forest* (0 vs. 1–3)**



**Fonte:** Elaboração própria (2026).

Por sua vez, o *Random Forest* apresenta maior especificidade e menor incidência de falsos positivos, evidenciando esse perfil mais conservador — característica desejável quando se busca reduzir alarmes indevidos em triagens clínicas.

Neste trabalho, no cenário binário, o termo conservador é empregado no sentido de maior especificidade (redução de falsos positivos), não devendo ser confundido com a estratégia conservadora baseada em risco adotada no cenário multiclasse. Adicionalmente, a consistência entre as curvas *ROC* e os valores de precisão–revocação sugere estabilidade probabilística dos modelos no regime de teste. Em síntese, os três modelos demonstraram capacidade robusta de discriminação binária, com diferenças sutis em seus perfis decisórios. A escolha do modelo mais adequado dependerá do regime operacional desejado — por exemplo, priorização da sensibilidade ou da precisão — e será discutida em maior profundidade no capítulo seguinte.

#### 4.1.2 Curva *ROC*

A avaliação complementar do cenário binário foi realizada por meio da curva *ROC*, a qual permite analisar a capacidade discriminativa global dos modelos ao longo de diferentes

limiares de decisão. Essa análise fornece uma visão consolidada da separação entre as classes, independentemente da escolha específica do limiar. A Tabela 2 sintetiza os valores de *AUC* obtidos no cenário binário (0 vs. 1–3), bem como a métrica *AP* associada à análise *Precision–Recall*, utilizada como medida complementar de desempenho.

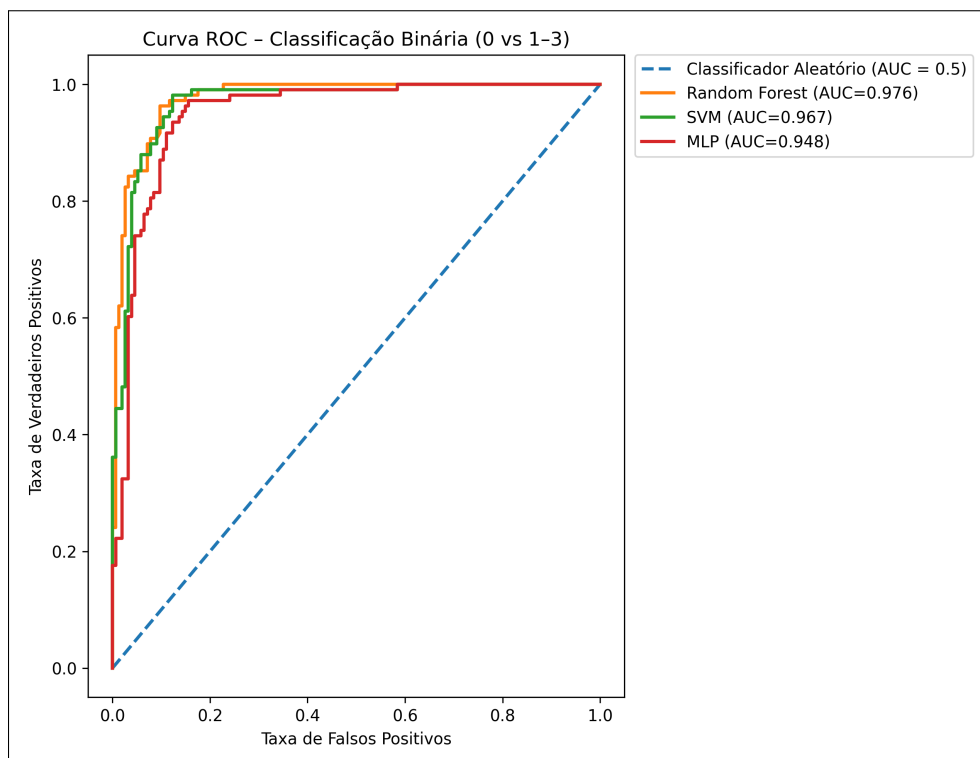
Tabela 2: *AUC (ROC) e AP (Precision–Recall) no cenário binário (0 vs 1–3).*

Modelo	<i>AUC (ROC) (%)</i>	<i>AP (PR) (%)</i>
<i>SVM</i>	96,70	95,00
<i>MLP</i>	94,80	93,20
<i>Random Forest</i>	97,60	96,10

Fonte: Elaboração própria (2026).

Observa-se que todas as arquiteturas apresentam elevada capacidade discriminativa, com curvas próximas ao canto superior esquerdo do plano e valores de *AUC* superiores a 0,94, indicando separação consistente entre as classes.

Figura 13 – Curvas *ROC* – Classificação Binária (0 vs. 1–3)



Fonte: Elaboração própria (2026).

Em síntese, os resultados evidenciam desempenho consistente dos modelos no cenário binário, sem a existência de um classificador universalmente superior. O *Random Forest* destaca-

se pela maior *AUC*, indicando maior estabilidade global, enquanto o *SVM* apresenta desempenho próximo. O *MLP*, embora ligeiramente inferior sob a métrica *ROC*, mantém capacidade discriminativa elevada, reforçando o caráter complementar das arquiteturas avaliadas no contexto de sistemas de triagem clínica.

## 4.2 Desempenho Multiclasse (0–3)

Neste cenário, os modelos foram avaliados na classificação em quatro níveis ordenados de risco (0–3). A análise multiclasse permite avaliar o desempenho por classe e a ocorrência de erros entre níveis adjacentes, aspecto relevante para aplicações clínicas. Os resultados do cenário multiclasse apresentados nesta seção refletem as previsões após a aplicação da estratégia de decisão conservadora descrita na Etapa 6 da metodologia, na qual casos de baixa confiança preditiva são direcionados para a classe de maior risco entre as duas classes mais prováveis, visando reduzir subestimações clinicamente relevantes.

Tabela 3: **Acertos por classe no cenário multiclasse (0–3).**

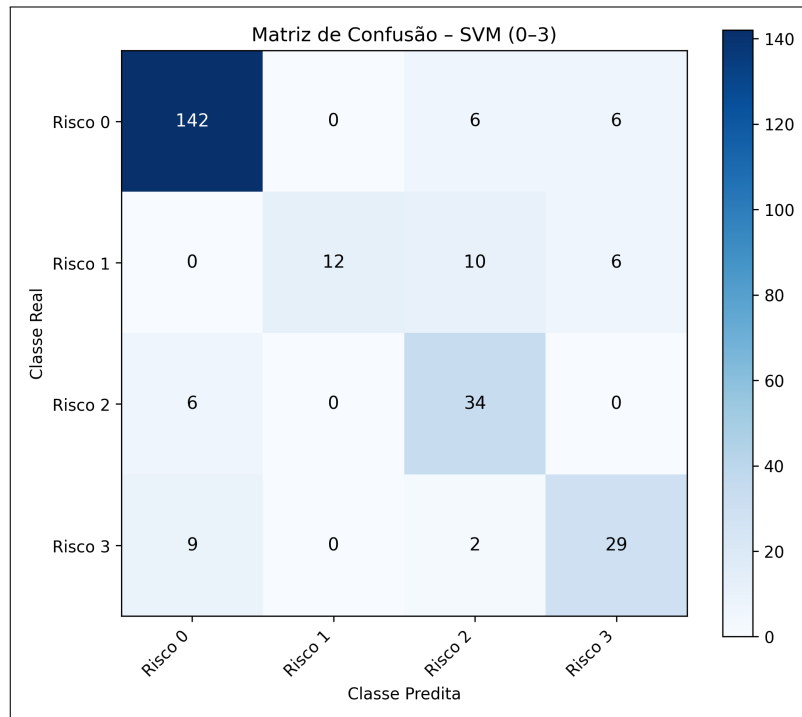
Modelo	Classe 0	Classe 1	Classe 2	Classe 3
<i>MLP</i>	139/154 (90,3%)	19/28 (67,9%)	33/40 (82,5%)	29/40 (72,5%)
<i>Random Forest</i>	146/154 (94,8%)	15/28 (53,6%)	24/40 (60,0%)	38/40 (95,0%)
<i>SVM</i>	142/154 (92,2%)	12/28 (42,9%)	34/40 (85,0%)	29/40 (72,5%)

**Nota:** Total de registros (crianças) por classe no conjunto de teste: Classe 0 = 154; Classe 1 = 28; Classe 2 = 40; Classe 3 = 40 ( $N_{\text{teste}} = 262$ ).

**Fonte:** Elaboração própria (2026).

No cenário multiclasse (0–3), os modelos apresentam maior acurácia nas classes extremas (0 e 3), enquanto as classes intermediárias (1 e 2) concentram a maior parte das confusões, refletindo a natureza gradual e ordinal do risco clínico. Pode-se observar que modelos apresentaram padrões de erro distintos, refletindo diferenças em seus perfis decisórios. O *MLP* apresentou confusões principalmente entre a classe 0 e as classes 2 e 3, além de dispersão da classe 1 entre níveis adjacentes de risco, indicando dificuldade na delimitação dos estados intermediários. O *Random Forest* evidenciou confusão relevante da classe 2 com a classe 0, bem como erros da classe 1 direcionados à classe 3, sugerindo uma tendência à superestimação do risco (atribuição a classes mais altas), comportamento que pode reduzir a probabilidade de subestimações clinicamente relevantes, embora possa aumentar a ocorrência de alarmes indevidos.

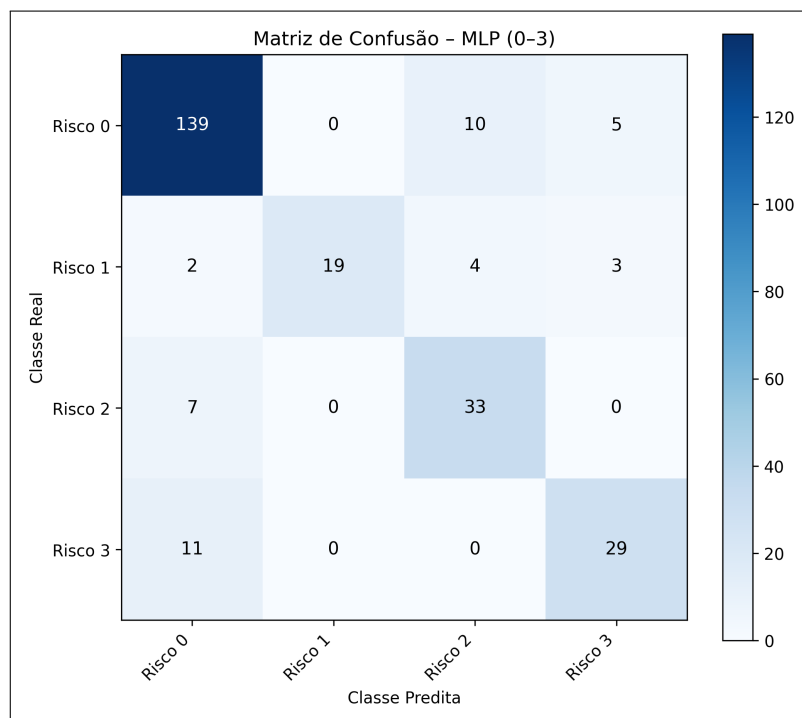
**Figura 14 – Matriz de Confusão Multiclasse do modelo SVM (0–3)**



Fonte: Elaboração própria (2026).

O SVM apresentou confusões frequentes da classe 1 com as classes 2 e 3, além de casos de classe 3 confundida com 0, indicando dificuldade na separação entre níveis intermediários e extremos — padrão típico em classificação ordinal.

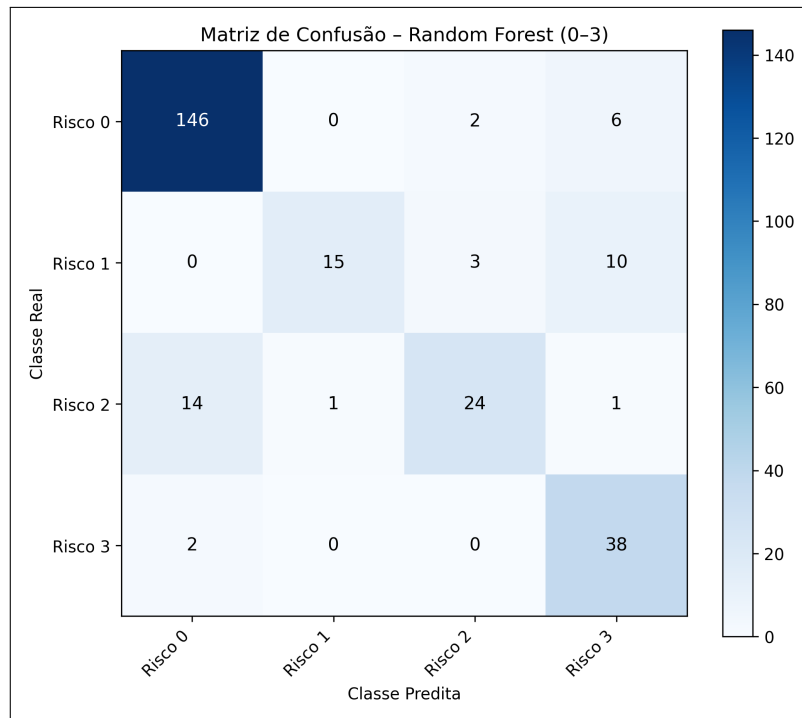
**Figura 15 – Matriz de Confusão Multiclasse do modelo MLP (0–3)**



Fonte: Elaboração própria (2026).

O *MLP* demonstra melhor retenção das classes intermediárias em relação aos demais modelos, especialmente na classe 1, sugerindo maior capacidade de modelar transições graduais entre níveis adjacentes de risco por meio de representações não lineares distribuídas.

**Figura 16 – Matriz de Confusão Multiclasse do modelo *Random Forest* (0–3)**



Fonte: Elaboração própria (2026).

O *Random Forest* apresenta desempenho particularmente elevado na classe 3, com alto número de acertos, porém exibe confusões relevantes entre as classes 0 e 2, indicando um comportamento mais rígido na definição de fronteiras entre níveis intermediários do risco.

Em síntese, os resultados multiclasse evidenciam que, embora os modelos sejam eficazes na identificação dos extremos do espectro de risco, a separação entre níveis intermediários permanece como o principal desafio. Essas confusões refletem tanto limitações estruturais dos dados quanto a própria natureza contínua e ordinal do construto clínico, aspecto que será aprofundado na análise discursiva do capítulo seguinte.

#### 4.2.1 Interpretabilidade e explicabilidade com *SHAP*

No contexto do cenário multiclasse (0–3), a interpretabilidade dos modelos de aprendizado de máquina constitui um requisito essencial, sobretudo em aplicações clínicas, nas quais é necessário compreender como os atributos de entrada influenciam as decisões do sistema. Nesse sentido, foi empregado o método *SHAP* (*SHapley Additive exPlanations*) para analisar o compor-

tamento do modelo *Multilayer Perceptron (MLP)*, permitindo identificar a contribuição relativa de cada variável na predição dos diferentes níveis de risco.

A Figura 17 apresenta a importância global das variáveis segundo o critério do impacto médio absoluto dos valores de Shapley. Esse tipo de visualização permite identificar quais atributos exercem maior influência nas decisões do modelo, considerando o efeito médio de cada variável ao longo de todas as amostras avaliadas. Observa-se que o *MLP* atribui maior relevância a um conjunto específico de variáveis, enquanto outras apresentam contribuição mais moderada ou residual.

**Figura 17 – Importância Global das Variáveis no Modelo *MLP* (*SHAP* Multiclasse)**



Fonte: Elaboração própria (2026).

Do ponto de vista metodológico, o *SHAP* baseia-se na teoria dos valores de Shapley, proveniente da teoria dos jogos cooperativos, na qual a contribuição total de um sistema é distribuída entre seus componentes de forma aditiva. Aplicado à modelagem computacional, cada variável de entrada é tratada como um agente que contribui para a saída final do modelo, permitindo decompor a predição em parcelas atribuíveis a cada atributo.

A análise das importâncias indica que o modelo utiliza múltiplas variáveis de forma conjunta, não se apoiando em um único atributo dominante. Esse comportamento sugere que as decisões do *MLP* emergem da combinação de diferentes dimensões dos dados psicométricos,

refletindo a natureza multidimensional do problema de classificação multiclasse. As variáveis de maior impacto concentram-se entre aquelas mais frequentemente associadas às respostas do modelo, enquanto atributos com menor relevância apresentam influência reduzida na decisão final.

Em termos de modelagem, esse padrão evidencia que o *MLP* constrói representações internas distribuídas, nas quais diferentes atributos contribuem simultaneamente para a predição do nível de risco. Essa característica é consistente com arquiteturas neurais não lineares, que integram múltiplas fontes de informação em camadas latentes antes da tomada de decisão.

Em síntese, a análise de interpretabilidade por meio do *SHAP* permite caracterizar o comportamento interno do modelo *MLP* no cenário multiclasse, identificando as variáveis mais influentes e evidenciando a natureza distribuída de suas decisões. Os resultados apresentados nesta subseção fornecem subsídios quantitativos e estruturais para a análise interpretativa, a qual será aprofundada no capítulo de Discussão.

#### 4.2.2 Erro clínico multiclasse

No cenário multiclasse (0–3), a avaliação do desempenho foi complementada por uma análise do erro clínico, considerando não apenas a frequência, mas também o deslocamento entre classes. Os resultados foram organizados em três categorias: **acertos**, **erros adjacentes** (atribuição à classe vizinha) e **erros não adjacentes** (deslocamentos entre classes não contíguas, associados a maior discrepância entre o nível real e o predito). A Tabela 4 apresenta os percentuais correspondentes a cada categoria.

Tabela 4: **Distribuição percentual do erro clínico no cenário multiclasse (0–3).**

Modelo	Acertos (%)	Erro adjacente (%)	Erro não adjacente (%)	Total (%)
<i>Random Forest</i>	85,1	1,9	13,0	100,0
<i>SVM</i>	82,8	4,6	12,6	100,0
<i>MLP</i>	84,0	2,3	13,7	100,0

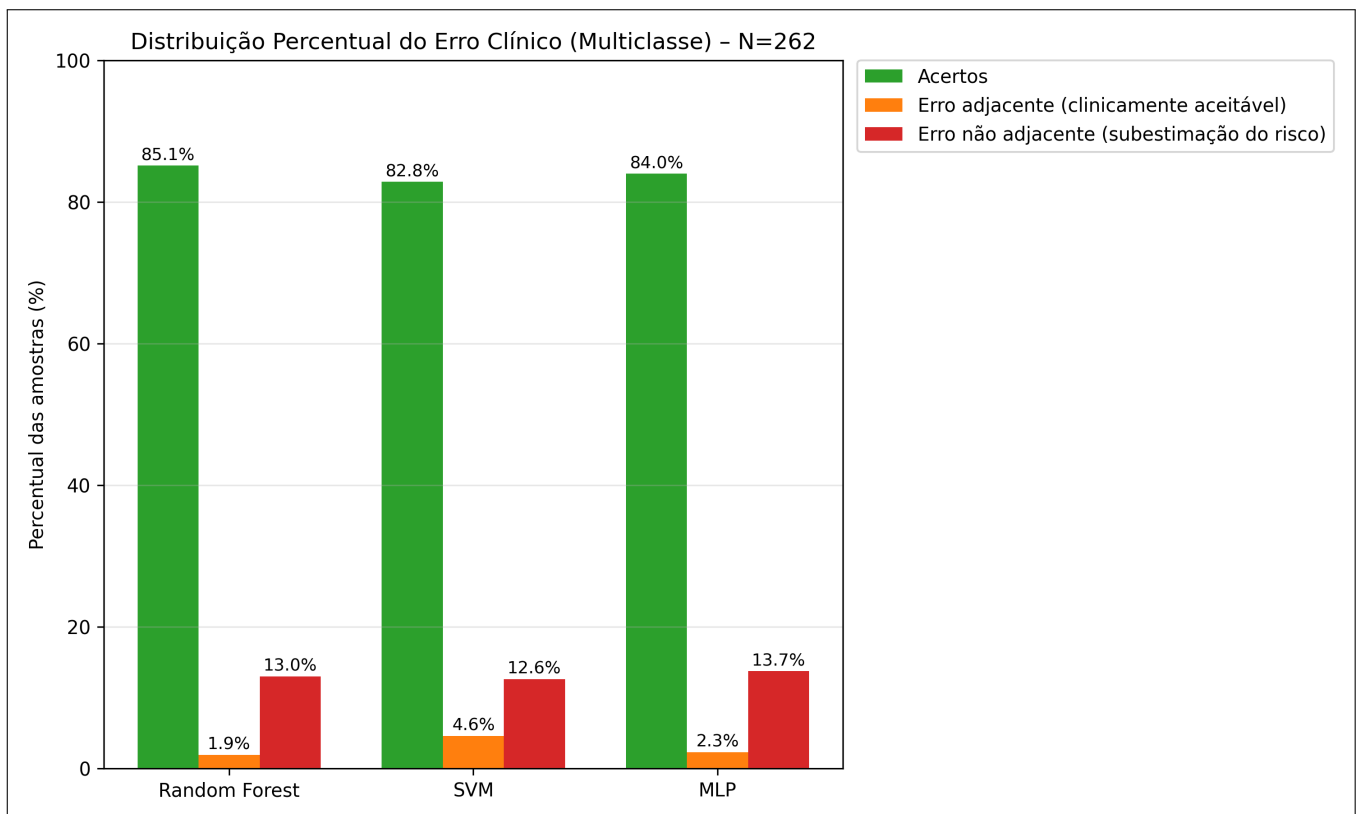
**Nota:** Erros adjacentes correspondem a confusões entre classes vizinhas; erros não adjacentes referem-se a deslocamentos entre classes não contíguas no espectro de risco.

**Fonte:** Elaboração própria (2026).

Os resultados indicam que o modelo *Random Forest* apresenta a maior taxa de acertos

globais (85,1%), acompanhada por baixa incidência de erros adjacentes (1,9%) e uma proporção de erros não adjacentes de 13,0%. O *SVM* alcança 82,8% de acertos, com maior frequência de erros adjacentes (4,6%) e 12,6% de erros não adjacentes. O *MLP* apresenta 84,0% de acertos, 2,3% de erros adjacentes e 13,7% de erros não adjacentes, refletindo um equilíbrio intermediário entre sensibilidade às classes adjacentes e severidade do erro. A Figura 18 representa a distribuição percentual dessas categorias para os modelos *Random Forest*, *Support Vector Machine (SVM)* e *Multilayer Perceptron (MLP)*, considerando um total de  $N_{\text{teste}} = 262$  registros (crianças) avaliados no conjunto de teste. Essa representação permite comparar os perfis decisórios dos classificadores no contexto multiclasse.

**Figura 18 – Distribuição Percentual do Erro Clínico (Multiclasse)**



Fonte: Elaboração própria (2026).

De forma geral, observa-se que os três modelos apresentam desempenho elevado no cenário multiclasse, porém com distribuições distintas entre erros adjacentes e erros não adjacentes. Esses resultados permitem caracterizar quantitativamente o comportamento dos classificadores em relação à severidade do erro, fornecendo subsídios objetivos para análises comparativas posteriores.

### 4.3 Síntese dos Resultados

Esta subseção consolida os principais resultados obtidos ao longo do processo experimental, integrando métricas de desempenho nos cenários binário e multiclasse, análise por classe, estabilidade de generalização e distribuição do erro clínico. De forma geral, o *Random Forest* apresentou o melhor desempenho global no cenário binário, enquanto no cenário multiclasse observou-se maior dificuldade nas classes intermediárias, com destaque do *MLP* nessas classes. Sob a perspectiva clínica, embora o *Random Forest* tenha exibido a maior taxa global de acertos, o *SVM* apresentou a menor proporção de erros não adjacentes, evidenciando compromissos distintos entre desempenho médio e severidade do erro. Assim, observa-se um compromisso entre desempenho médio global (acertos) e severidade do erro (erros não adjacentes), aspecto relevante na interpretação clínica e que será aprofundado no capítulo de Discussão.

#### 4.3.1 Cenário binário (0 vs 1–3)

A avaliação comparativa no cenário binário foi conduzida com base em métricas clássicas de desempenho, permitindo uma análise quantitativa direta entre os modelos avaliados.

A Tabela 5 apresenta a síntese das métricas obtidas, incluindo acurácia, precisão, sensibilidade, especificidade e área sob a curva *ROC (AUC)*.

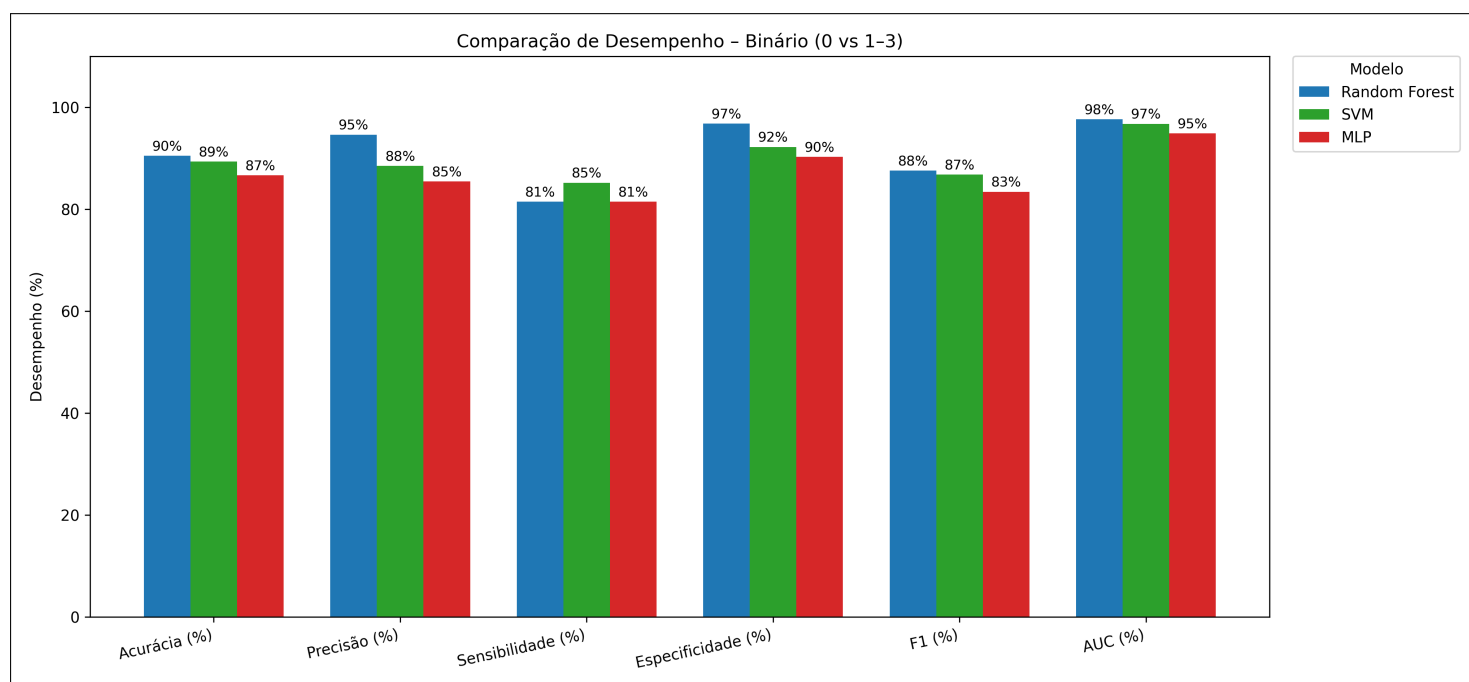
Tabela 5: Síntese das métricas de desempenho no cenário binário (0 vs 1–3).

Modelo	Acurácia (%)	Precisão (%)	Sensibilidade (%)	Especificidade (%)	AUC (%)
<i>Random Forest</i>	90,46	94,62	81,48	96,75	97,60
<i>MLP</i>	86,64	85,44	81,48	90,26	94,80
<i>SVM</i>	89,31	88,46	85,19	92,21	96,70

Fonte: Elaboração própria (2026).

Os resultados indicam superioridade do *Random Forest* no cenário binário, com os maiores valores de acurácia, precisão, especificidade e *AUC*, refletindo elevada capacidade discriminativa para a tarefa de triagem clínica. A Figura 19 complementa a análise tabular por meio de uma representação gráfica comparativa das métricas, facilitando a visualização das diferenças relativas entre os classificadores.

**Figura 19 – Comparação de desempenho no cenário binário (0 vs 1–3)**



Fonte: Elaboração própria (2026).

#### 4.3.2 Cenário multiclasse (0–3)

A avaliação no cenário multiclasse teve como objetivo caracterizar o comportamento dos modelos ao longo dos diferentes níveis de risco, permitindo identificar variações específicas de desempenho por classe. A Tabela 6 apresenta os valores de *F1-score* por classe, métrica que reflete o equilíbrio entre precisão e sensibilidade para cada nível de risco individualmente.

Tabela 6: *F1-score* por classe no cenário multiclasse (0–3).

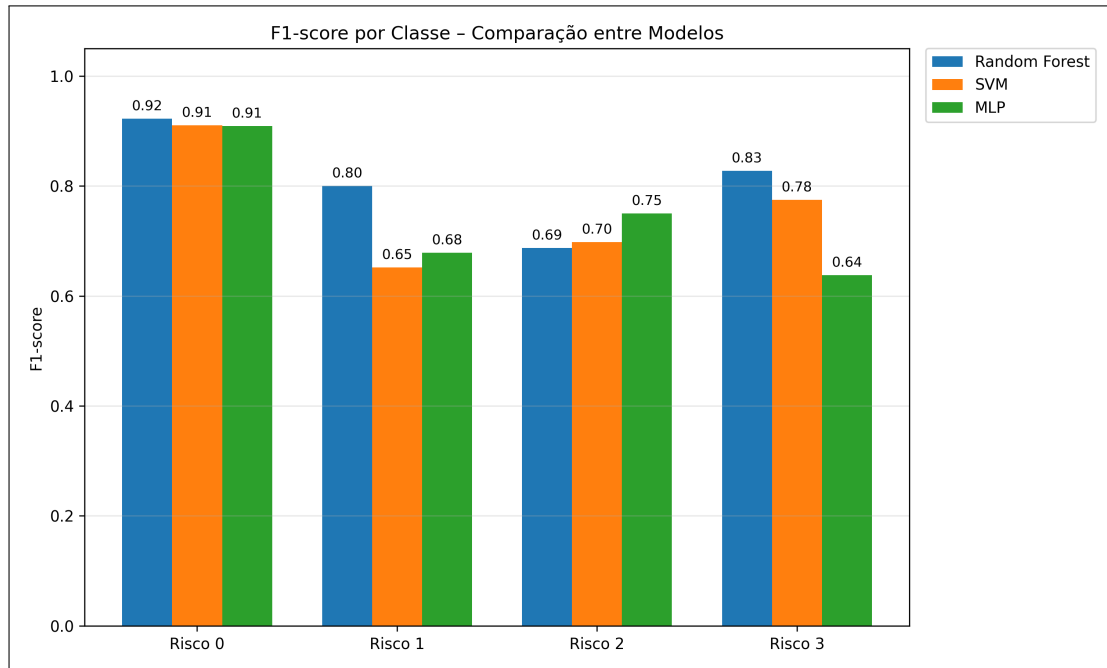
Modelo	Classe 0	Classe 1	Classe 2	Classe 3
<i>Random Forest</i>	0,94	0,61	0,67	0,83
<i>MLP</i>	0,93	0,63	0,71	0,79
<i>SVM</i>	0,92	0,57	0,69	0,75

Fonte: Elaboração própria (2026).

Observa-se que todos os modelos apresentam desempenho elevado na classe 0 (ausência de risco), enquanto as classes intermediárias, especialmente a classe 1, concentram os menores valores de *F1-score*, indicando maior dificuldade discriminativa. Destaca-se o *MLP* nas classes 1 e 2, sugerindo maior capacidade de modelar transições graduais entre níveis adjacentes de

risco. A Figura 20 apresenta a visualização comparativa desses resultados.

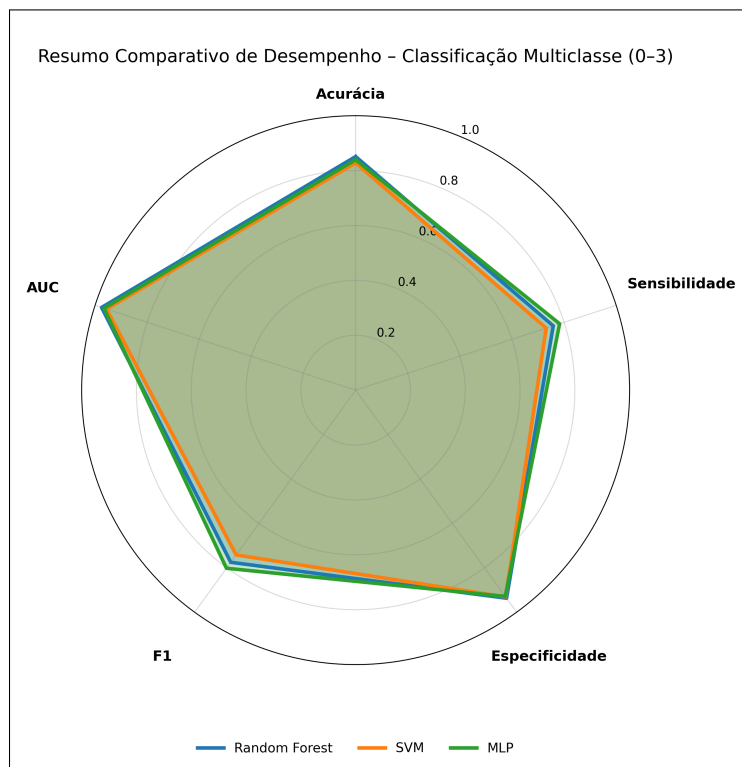
**Figura 20 – F1-score por classe no cenário multiclasse (0–3)**



Fonte: Elaboração própria (2026).

A Figura 21 sintetiza o desempenho global dos modelos no cenário multiclasse por meio de um gráfico radar, permitindo uma comparação integrada entre múltiplas métricas.

**Figura 21 – Resumo comparativo multiclasse por gráfico radar**



Fonte: Elaboração própria (2026).

### 4.3.3 Erro clínico no cenário multiclasse

A análise do erro clínico teve como objetivo qualificar a gravidade das decisões incorretas no cenário multiclasse, distinguindo entre classificações corretas, erros adjacentes — correspondentes a confusões entre classes vizinhas no espectro de risco — e erros não adjacentes, caracterizados por deslocamentos entre classes não contíguas. A distribuição percentual dessas categorias para cada modelo foi apresentada anteriormente na Tabela 4, permitindo uma avaliação comparativa direta entre os classificadores.

De modo geral, observa-se que todos os modelos apresentam elevada taxa de acertos globais, superiores a 80%, evidenciando desempenho consistente mesmo no cenário multiclasse, que representa a configuração mais complexa deste estudo. O modelo *Random Forest* alcançou o maior percentual de acertos (85,1%), acompanhado por baixa incidência de erros adjacentes (1,9%) e proporção moderada de erros não adjacentes (13,0%), indicando um perfil decisório mais estável. O *SVM* apresentou 82,8% de acertos, com maior frequência de erros adjacentes (4,6%), porém a menor taxa de erros não adjacentes (12,6%), sugerindo maior segurança clínica em termos de menor ocorrência de erros não adjacentes (deslocamentos amplos no espectro de risco). O *MLP*, por sua vez, obteve 84,0% de acertos, com 2,3% de erros adjacentes e 13,7% de erros não adjacentes, refletindo um equilíbrio intermediário entre sensibilidade às classes adjacentes e severidade do erro.

Observa-se ainda que os erros adjacentes ocorreram em proporção inferior aos erros não adjacentes, sugerindo que parte das decisões incorretas envolveu deslocamentos mais amplos no espectro de risco. Esse comportamento sugere que, na ausência de otimização explícita para perda ordinal, os limiares decisórios dos classificadores podem induzir deslocamentos mais amplos no espectro de risco.

Em síntese, a análise do erro clínico evidencia que, embora existam diferenças nos perfis decisórios das arquiteturas investigadas, todos os modelos apresentam desempenho elevado no cenário multiclasse, com compromissos claros entre taxa global de acertos, sensibilidade às classes intermediárias e gravidade do erro clínico. Esses resultados fornecem subsídios objetivos para a discussão sobre a adequação de cada modelo como núcleo de um sistema de apoio à decisão clínica, a ser aprofundada no capítulo seguinte.

## 4.4 Discussão dos Resultados

Os resultados apresentados nas seções anteriores permitem uma análise aprofundada do desempenho dos modelos supervisionados *Random Forest*, *Support Vector Machine (SVM)* e *Multilayer Perceptron (MLP)* na tarefa de classificação do risco de transtornos de ansiedade pediátrica, tanto no cenário binário quanto no cenário multiclasse. A discussão a seguir interpreta esses achados à luz da literatura científica recente e dos referenciais conceituais apresentados no Capítulo de Referencial Teórico, com ênfase nas implicações metodológicas, clínicas e computacionais sob a perspectiva da Engenharia Elétrica aplicada à saúde.

No cenário binário (0 vs 1–3), os três modelos avaliados apresentaram desempenho elevado, com valores de *AUC* entre 0,948 e 0,976 e acurácias acima de 86%, indicando elevada capacidade discriminativa entre ausência e presença de risco. Esses resultados são consistentes com estudos prévios no campo da psiquiatria computacional, nos quais algoritmos supervisionados clássicos, como *Random Forest* e *SVM*, demonstraram robustez na identificação de estados clínicos latentes a partir de dados psicométricos e comportamentais estruturados, conforme discutido por Shatte et al. [6] e Rivera et al. [22]. Em particular, o *Random Forest* apresentou o melhor desempenho global nesse cenário, refletido por maiores valores de *AUC* e menor incidência de falsos positivos, o que caracteriza um perfil mais conservador no sentido de maior especificidade (redução de FP) — propriedade desejável em aplicações de triagem clínica quando se busca reduzir alarmes indevidos.

A análise das matrizes de confusão binárias evidencia diferenças sutis, porém relevantes, nos perfis decisórios dos classificadores. Enquanto o *Random Forest* apresentou maior especificidade, o *SVM* exibiu equilíbrio mais pronunciado entre sensibilidade e especificidade, e o *MLP* manteve desempenho consistente em ambos os critérios. Esse comportamento está alinhado às propriedades teóricas dessas arquiteturas: modelos baseados em *ensembles*, como o *Random Forest*, tendem a reduzir variância e favorecer decisões mais estáveis, conforme proposto por Breiman [13], enquanto métodos de margem máxima, como o *SVM*, apresentam boa generalização em bases de dimensão moderada e com ruído estrutural, características típicas de dados clínicos, segundo Cortes e Vapnik [12].

No cenário multiclasse (0–3), o problema torna-se substancialmente mais complexo, refletindo a natureza contínua e ordinal do risco de ansiedade pediátrica. Os resultados mostram que todos os modelos apresentaram elevado desempenho nas classes extremas — classe 0 (ausência de risco) e classe 3 (risco elevado) — e maior dificuldade na separação das classes inter-

mediárias (1 e 2). Esse padrão é amplamente reportado na literatura e decorre da sobreposição fenotípica entre níveis adjacentes de risco, bem como da própria natureza subjetiva das medidas psicométricas utilizadas, conforme discutido por Kraemer [39] e Allen et al. [69]. Assim, a redução do *F1-score* nas classes intermediárias não deve ser interpretada como falha do modelo, mas como reflexo das limitações estruturais do processo de observação e rotulagem clínica.

A análise do *F1-score* por classe evidencia que diferentes modelos se destacam em diferentes regiões do espectro de risco. O *MLP* apresentou melhor desempenho na classe 1 e na classe 2, enquanto o *Random Forest* obteve maior desempenho na classe 3, e o *SVM* apresentou desempenho consistente, porém com maior dificuldade na classe 1. Esse comportamento sugere que as arquiteturas capturam padrões distintos nos dados, reforçando a ideia de complementaridade entre os modelos. Em particular, a capacidade do *MLP* de modelar relações não lineares distribuídas pode favorecer a identificação de padrões sutis associados a níveis iniciais e intermediários de risco, enquanto a estrutura de decisão do *Random Forest* pode favorecer a identificação de padrões mais consistentes associados a quadros de maior gravidade.

A utilização do gráfico radar como síntese multivariada permite observar o equilíbrio global entre métricas de desempenho no cenário multiclasse, evidenciando que nenhum modelo domina completamente os demais em todos os critérios avaliados. Essa visualização reforça a noção de que a escolha do classificador mais adequado depende do regime operacional desejado, bem como do custo clínico associado aos diferentes tipos de erro. Tal constatação está em consonância com abordagens contemporâneas de decisão em saúde, que defendem a adoção de modelos ajustados ao contexto clínico específico, em vez da busca por um classificador universalmente ótimo, conforme discutido por Talaei-Khoei et al. [18].

A análise do erro clínico multiclasse constitui um dos principais diferenciais deste trabalho em relação ao estado da arte. Ao distinguir entre acertos, erros adjacentes e erros não adjacentes, a avaliação transcende métricas puramente estatísticas e incorpora explicitamente a gravidade clínica do erro de decisão. Os resultados indicam que os três modelos apresentam taxas de acerto superiores a 80% no conjunto de teste ( $N_{\text{teste}} = 262$ ), evidenciando desempenho global consistente. Observa-se, entretanto, que os erros não adjacentes apresentaram percentuais superiores aos erros adjacentes. Esse comportamento pode refletir a sobreposição estrutural entre extremos do espectro de risco em dados psicométricos ruidosos, a presença de limiares decisórios rígidos nas arquiteturas supervisionadas e, em parte, a estratégia conservadora adotada, que prioriza evitar subestimações clínicas mesmo ao custo de deslocamentos mais amplos entre

níveis de risco. Cabe observar que essa estratégia foi aplicada de forma uniforme aos três modelos, de modo que as diferenças observadas refletem propriedades intrínsecas das arquiteturas, e não viés procedimental. Esse padrão sugere que, embora o sistema discrimine adequadamente grande parte dos casos, uma parcela das decisões incorretas envolve deslocamentos mais amplos entre níveis de risco, possivelmente associados à maior complexidade, à sobreposição de perfis clínicos e à incerteza de rótulo nas classes intermediárias, conforme discutido por Kraemer [39]. Ainda assim, os percentuais de erro não adjacente mantiveram-se controlados (aproximadamente 12–14%), significativamente inferiores à taxa global de acerto, o que sustenta a viabilidade do sistema como ferramenta de apoio à triagem.

Cabe destacar que, no contexto multiclasse, o *Random Forest* apresentou indícios de superestimação do risco (atribuição a classes mais altas), comportamento que pode reduzir subestimações clinicamente graves, porém aumentar alarmes indevidos. Nesse cenário, o termo “conservador” não se refere à redução de falsos positivos (como no binário), mas à priorização da segurança clínica ao evitar subestimações severas. Já o *SVM* apresentou a menor proporção de erros não adjacentes, sugerindo maior segurança em termos de menor ocorrência de deslocamentos amplos no espectro ordinal. O *MLP*, por sua vez, apresentou equilíbrio intermediário, com melhor retenção das classes intermediárias, embora com proporção de erros não adjacentes semelhante aos demais.

Sob a perspectiva da Engenharia Elétrica, esses achados podem ser interpretados como indicativos de que o sistema de decisão opera em um regime de separabilidade limitada pelo ruído, pela sobreposição estrutural das classes e pela incerteza de rotulagem, e não necessariamente por deficiência do modelo computacional. Esse cenário é análogo a problemas clássicos de detecção de sinais em ambientes ruidosos, nos quais o desempenho do classificador é naturalmente limitado pela relação sinal–ruído do processo de medição.

A análise de interpretabilidade baseada em *SHAP* complementa a avaliação quantitativa ao evidenciar que o *MLP* constrói suas decisões a partir da combinação de múltiplas variáveis psicométricas, sem dependência excessiva de um único atributo dominante. Esse comportamento é desejável em aplicações clínicas, pois sugere que o modelo capta a natureza multidimensional do construto ansiedade, em consonância com os referenciais diagnósticos do DSM-5 [10], e da CID-11 [11]. Além disso, a coerência entre as variáveis mais influentes e os construtos clínicos reconhecidos reforça a plausibilidade clínica do modelo e sua potencial utilidade como ferramenta de apoio à decisão.

Em termos de contribuição científica, este trabalho avança o estado da arte ao demonstrar que a classificação do risco de ansiedade pediátrica pode ser tratada de forma mais realista e clinicamente informativa quando modelada como um problema de decisão ordinal, em vez de uma tarefa puramente binária. Enquanto grande parte da literatura em psiquiatria computacional concentra-se na distinção presença/ausência de transtorno, os resultados aqui apresentados mostram que a estratificação em múltiplos níveis de risco fornece informações adicionais relevantes para a prática clínica, especialmente no contexto de triagem e priorização de encaminhamentos.

Do ponto de vista computacional, os resultados evidenciam que arquiteturas supervisionadas clássicas — *Random Forest*, *Support Vector Machine* e *Multilayer Perceptron* — apresentam desempenhos comparáveis em termos globais, porém com perfis decisórios distintos. No cenário binário, o *Random Forest* mostrou-se mais específico (menor FP), enquanto o *SVM* apresentou maior equilíbrio entre sensibilidade e especificidade, e o *MLP* manteve desempenho consistente. No cenário multiclasse, observa-se maior dificuldade nas classes intermediárias, com destaque do *MLP* nessas classes, ao passo que o *Random Forest* se destacou na classe 3 e o *SVM* apresentou menor proporção de erros não adjacentes. Essa diferenciação reforça a ideia de que não existe um classificador universalmente ótimo, mas sim modelos mais adequados a regimes operacionais específicos, conceito alinhado a abordagens modernas de engenharia de sistemas em saúde, conforme discutido por Steyerberg [70].

A opção por arquiteturas supervisionadas clássicas, em detrimento de modelos mais recentes baseados em *deep learning*, foi uma decisão metodológica deliberada, fundamentada nas características do problema e do conjunto de dados. Em particular, o tamanho amostral moderado, a natureza tabular e psicométrica das variáveis, bem como a necessidade de interpretabilidade clínica, tornam modelos profundos menos adequados e potencialmente mais suscetíveis a sobreajuste. Estudos recentes, como os de Shatte et al. [6] e Grinsztajn et al. [71], demonstram que, em cenários com dados estruturados e número limitado de amostras, algoritmos clássicos frequentemente igualam ou superam o desempenho de redes profundas, com maior estabilidade e menor custo computacional. Além disso, arquiteturas profundas tendem a operar como modelos de alta capacidade com baixa transparência decisória, o que representa uma limitação relevante em aplicações clínicas sensíveis, nas quais a rastreabilidade da decisão e a compreensão do erro são fundamentais para a aceitação do sistema. Assim, a escolha por modelos clássicos favorece não apenas a robustez estatística, mas também a confiabilidade, a interpretabilidade e a viabilidade de implementação em sistemas de apoio à decisão clínica, em consonância com reco-

mendações contemporâneas para o uso responsável de inteligência artificial em saúde, conforme discutido por Topol [72] e Holzinger et al. [73].

Sob a ótica da Medicina e da Psiquiatria do Desenvolvimento, os achados corroboram a concepção dimensional dos transtornos de ansiedade, conforme discutido por Kraemer [39] e adotado em classificações diagnósticas contemporâneas, como o DSM-5 [10] e CID-11 [11]. A maior taxa de acerto nas classes extremas (ausência de risco e risco elevado) e a maior ambiguidade nas classes intermediárias refletem a própria natureza gradual da manifestação sintomática, especialmente na primeira infância, em que limites diagnósticos são menos nítidos e altamente dependentes do contexto familiar e do desenvolvimento neuropsicológico.

Em termos de Saúde Pública, os resultados indicam que sistemas baseados em aprendizado de máquina podem contribuir para mitigar a lacuna histórica entre prevalência e acesso ao cuidado em saúde mental infantil. Conforme a Organização Mundial da Saúde e o UNICEF [74], cerca de 1 em cada 7 crianças e adolescentes apresenta condições de saúde mental, com os transtornos de ansiedade figurando entre os mais prevalentes. Ademais, Wang et al. [75] demonstram, por meio de revisão sistemática e meta-análise recente, que apenas aproximadamente 31% das crianças e adolescentes com transtornos mentais recebem tratamento adequado. Esse cenário evidencia a subutilização dos serviços especializados e reforça a relevância de ferramentas computacionais capazes de estratificar risco de forma automatizada, interpretável e escalável, auxiliando na priorização de encaminhamentos e na otimização de recursos na atenção primária.

Quando comparado a trabalhos de referência que utilizaram a mesma base de dados, como o estudo seminal de Carpenter et al. [9], este trabalho apresenta uma abordagem complementar. Enquanto Carpenter et al. [9] focaram na redução dimensional e na identificação de subconjuntos mínimos de itens do *Preschool Age Psychiatric Assessment (PAPA)* — instrumento diagnóstico estruturado amplamente utilizado para avaliação psiquiátrica em crianças em idade pré-escolar — para rastreamento de transtornos específicos (TAG e TAS), com acurácias superiores a 96%, o presente estudo amplia o escopo ao investigar a estratificação ordinal de risco (0–3) e a severidade do erro clínico. Essa diferença de enfoque não configura competição direta, mas sim extensão metodológica, aproximando a modelagem do raciocínio clínico contínuo utilizado na prática.

Por fim, a incorporação de técnicas de interpretabilidade, como a análise baseada em *SHAP*, fortalece a plausibilidade clínica dos modelos ao evidenciar coerência entre as variáveis mais influentes e os construtos clínicos reconhecidos na literatura. Tal característica é particular-

mente relevante diante das preocupações éticas e metodológicas associadas ao uso de inteligência artificial em saúde mental, conforme discutido por Bone et al. [76] e por revisões sistemáticas mais recentes, como a de Talaei-Khoei et al. [18]. Dessa forma, o trabalho não apenas demonstra desempenho quantitativo robusto, mas também contribui para o debate sobre transparência, segurança e aplicabilidade clínica de sistemas inteligentes em saúde mental infantil.

O quadro comparativo apresentado na Figura 22 permite situar os resultados obtidos neste estudo no contexto da literatura recente sobre aprendizado de máquina aplicado à ansiedade infantil e à saúde mental. Observa-se, inicialmente, que o desempenho binário alcançado (*AUC* entre 0,948 e 0,976) posiciona os modelos acima da faixa típica reportada em revisões sistemáticas da área, que frequentemente apresentam valores entre 0,75 e 0,90 para tarefas diagnósticas em saúde mental, conforme relatado por Shatte et al. [6]. Esse resultado indica elevada capacidade discriminativa na tarefa de triagem entre ausência e presença de risco, mantendo competitividade frente a estudos consolidados.

**Figura 22 – Resultados comparativos com estudos anteriores**

Estudo	Tarefa	N	Modelos	Desempenho Reportado
<b>Este trabalho (2026)</b>	Binária (0 vs 1–3) e Multiclasse (0–3)	262	RF, SVM, MLP	<i>AUC</i> = 0,94 (binário); Acurácia multiclasse ≈ 83–85%; erro não adjacente ≈ 12–14%
<b>Carpenter et al. (2016)</b>	Binária (TAG/TAS)	≈ 400	ADTree	Acurácia > 96%; Sensibilidade > 95%; modelo reduzido com ~8–10 itens
<b>Allen et al. (2019)</b>	Classificação dimensional	≈ 300	SVM, RF	Acurácia 75–82%; queda de desempenho em classes intermediárias
<b>Shatte et al. (2019)</b>	Revisão sistemática	--	DT, RF, SVM, Naive Bayes, MLP, KNN	Acurácia média em saúde mental: 70–88%; <i>AUC</i> típica: 0,75–0,90
<b>Talaei-Khoei et al. (2023)</b>	IA explicável (saúde)	--	RF, XGBoost, SVM, CNN, RNN, MLP	<i>AUC</i> relatadas entre 0,80–0,92; enfatizam análise do impacto clínico do erro

Fonte: Elaboração própria (2026).

Quando comparado ao trabalho de Carpenter et al. [9], com acurácia superior a 96% em tarefa binária específica para TAG/TAS, sob protocolo experimental distinto e com subconjunto reduzido de variáveis, é importante destacar diferenças metodológicas substanciais.

Enquanto aquele estudo foca na identificação de transtornos isolados e utiliza uma estrutura reduzida de variáveis (aproximadamente 8 a 10 itens), o presente trabalho explora um espectro ordinal completo de risco (0–3), incorporando maior heterogeneidade clínica e complexidade decisória. Dessa forma, embora os valores absolutos de acurácia sejam inferiores, a tarefa aqui proposta apresenta maior grau de dificuldade estrutural e maior proximidade com cenários reais de triagem graduada.

A acurácia global no cenário multiclasse foi de 85,1% para o Random Forest, 84,0% para o MLP e 82,8% para o SVM. Isso revela desempenho consistente diante de uma tarefa ordinal naturalmente mais desafiadora. Estudos como o de Allen et al. [69] também relatam redução de desempenho em abordagens dimensionais, especialmente nas classes intermediárias, fenômeno igualmente observado neste trabalho. Tal comportamento sugere que a ambiguidade entre níveis adjacentes de risco constitui uma característica intrínseca dos dados psicométricos pediátricos, e não necessariamente uma limitação específica do modelo adotado.

Outro aspecto relevante evidenciado na comparação refere-se à análise do erro clínico ordinal. Diferentemente da maior parte dos estudos comparados, que se concentram exclusivamente em métricas agregadas (acurácia, *AUC*, *FI*), este trabalho introduz a distinção explícita entre erros adjacentes e não adjacentes. A taxa de erro não adjacente entre aproximadamente 12% e 14% no cenário multiclasse indica que, apesar de existir uma parcela de decisões incorretas com deslocamentos amplos, sua ocorrência permanece significativamente menor do que o volume de classificações corretas, reforçando a viabilidade do sistema como ferramenta de triagem e apoio à decisão. Essa abordagem está alinhada às recomendações contemporâneas de avaliação centrada no impacto da decisão, conforme discutido por Talaei-Khoei et al. [18], ampliando a interpretação do desempenho para além de indicadores puramente estatísticos.

Em síntese, a comparação quantitativa indica que os modelos supervisionados clássicos avaliados apresentam desempenho competitivo em relação ao estado da arte, especialmente considerando a maior complexidade da tarefa ordinal proposta. Mais do que maximizar métricas globais, o presente trabalho contribui metodologicamente ao integrar análise de erro clínico ordinal e interpretabilidade estruturada, oferecendo uma avaliação mais alinhada às exigências de segurança, transparência e aplicabilidade clínica em sistemas de apoio à decisão para ansiedade pediátrica.

## 5 CONCLUSÃO E PERSPECTIVAS FUTURAS

Este trabalho apresentou o desenvolvimento e a avaliação de um sistema computacional de apoio à decisão para a classificação do risco de transtornos de ansiedade em crianças em idade pré-escolar, considerando tanto o cenário binário (ausência vs. presença de risco) quanto o cenário multiclasse ordinal (níveis 0–3). A abordagem foi fundamentada na perspectiva da Engenharia Elétrica aplicada à saúde mental, tratando o problema como um processo de inferência de estados latentes a partir de observações psicométricas estruturadas.

Com base na base pública disponibilizada pelo *Harvard Dataverse*, foram implementados e comparados três modelos supervisionados amplamente consolidados na literatura — *Random Forest*, *Support Vector Machine (SVM)* e *Multilayer Perceptron (MLP)* — avaliados por meio de validação cruzada estratificada, métricas clássicas de desempenho e análise explícita da gravidade do erro clínico. A metodologia integrou avaliação por *AUC*, precisão, sensibilidade, especificidade, *F1-score* por classe, bem como distinção entre erros adjacentes e não adjacentes no cenário multiclasse, além de análise de interpretabilidade baseada em *SHAP*.

No cenário binário (0 vs 1–3), todos os modelos apresentaram elevada capacidade discriminativa, com valores de *AUC* superiores a 0,94 e acurácias próximas ou superiores a 89%. O *Random Forest* obteve o melhor desempenho global, com maior acurácia e maior especificidade, evidenciando perfil mais conservador no sentido estatístico — isto é, menor incidência de falsos positivos. O *SVM* apresentou equilíbrio consistente entre sensibilidade e especificidade, enquanto o *MLP* demonstrou desempenho estável e competitivo, confirmando a adequação de arquiteturas neurais rasas à modelagem de dados psicométricos tabulares.

No cenário multiclasse (0–3), a tarefa revelou-se substancialmente mais desafiadora, refletindo a natureza ordinal e contínua do risco clínico. Os modelos apresentaram maior desempenho nas classes extremas (0 e 3) e maior dificuldade nas classes intermediárias (1 e 2), padrão compatível com a sobreposição fenotípica característica de construtos psicométricos em saúde mental infantil. As acurácias globais situaram-se entre aproximadamente 83% e 85%, demonstrando desempenho consistente mesmo diante da maior complexidade decisória.

A análise do erro clínico ordinal constitui um dos principais diferenciais metodológicos deste trabalho. Ao distinguir entre acertos, erros adjacentes (confusões entre classes vizinhas) e erros não adjacentes (deslocamentos mais amplos no espectro de risco), a avaliação incorporou explicitamente a gravidade clínica da decisão incorreta. Observou-se que os percentuais de erro não adjacente permaneceram entre aproximadamente 12% e 14%, enquanto as taxas glo-

bais de acerto superaram 80% para todos os modelos. Embora os erros não adjacentes exijam atenção sob a perspectiva clínica, sua ocorrência manteve-se significativamente inferior ao volume de classificações corretas, sustentando a viabilidade do sistema como ferramenta de apoio à triagem.

A análise de interpretabilidade por meio do método *SHAP* evidenciou que o modelo *MLP* constrói suas decisões a partir da combinação distribuída de múltiplas variáveis psicométricas, sem dependência excessiva de um único atributo dominante. Esse comportamento é coerente com a natureza multidimensional dos transtornos de ansiedade pediátrica e reforça a plausibilidade clínica do modelo, além de atender a requisitos éticos e metodológicos associados à explicabilidade em sistemas de inteligência artificial aplicados à saúde.

Sob a ótica da Engenharia Elétrica, os resultados sugerem que o limite de desempenho observado está mais associado à sobreposição estrutural das classes e à incerteza inerente aos rótulos clínicos do que à limitação dos algoritmos supervisionados empregados. O problema pode ser interpretado como um sistema de detecção operando sob relação sinal–ruído limitada, em que a separabilidade entre estados latentes é parcialmente restringida pela natureza do processo de medição psicométrica.

De forma geral, os achados demonstram que modelos supervisionados clássicos, quando corretamente validados, interpretados e contextualizados clinicamente, constituem ferramentas promissoras para a estratificação do risco de ansiedade pediátrica, especialmente em cenários de triagem e priorização de encaminhamentos.

## 5.1 Limitações do Estudo

Algumas limitações devem ser consideradas na interpretação dos resultados. A base de dados utilizada é composta exclusivamente por variáveis psicométricas estruturadas, não incorporando sinais psicofisiológicos ou neurobiológicos contínuos, como EEG ou variabilidade da frequência cardíaca, que poderiam ampliar o espaço observacional e potencialmente melhorar a separabilidade entre classes.

Além disso, os rótulos clínicos utilizados como *ground truth* baseiam-se em critérios diagnósticos derivados do DSM-5 [10] e CID-11 [11], os quais envolvem variabilidade interobservador e incerteza inerente à avaliação subjetiva. Tal característica impõe um limite superior teórico ao desempenho dos modelos supervisionados, especialmente no cenário multiclasse.

Outra limitação relevante refere-se à ausência de validação externa em bases indepen-

dentes, restringindo a generalização populacional dos achados. Embora a validação cruzada estratificada reduza o risco de sobreajuste, estudos futuros com coortes distintas são necessários para avaliar a robustez do sistema em diferentes contextos clínicos e socioculturais.

## 5.2 Perspectivas Futuras

Como perspectivas futuras, destaca-se a exploração de modelos explicitamente ordinais ou sensíveis a custos assimétricos de erro, com funções de perda customizadas e ajuste de limites decisórios, visando reduzir ainda mais a probabilidade de subestimação clínica.

Outra direção promissora envolve a integração de dados multimodais, combinando informações psicométricas com sinais psicofisiológicos e neurobiológicos. Sob a perspectiva da Engenharia Elétrica, essa ampliação pode melhorar a relação sinal–ruído do sistema de inferência e favorecer a discriminação entre níveis intermediários de risco.

Adicionalmente, investigações futuras podem contemplar estratégias de aprendizado semi-supervisionado ou técnicas de calibração probabilística mais refinadas, especialmente em cenários clínicos com limitação de rótulos confiáveis. A ampliação das técnicas de explicabilidade, incluindo análises locais individualizadas, também representa avanço relevante para fortalecer a confiança clínica e a aceitação institucional do sistema.

Por fim, a aplicação do *pipeline* desenvolvido em ambientes clínicos reais e sua adaptação para sistemas de triagem automatizada em larga escala configuram perspectivas de impacto direto em saúde pública. A consolidação de ferramentas computacionais objetivas, interpretáveis e alinhadas aos referenciais diagnósticos internacionais representa passo importante para intervenções precoces, priorização eficiente de recursos e suporte quantitativo à tomada de decisão em saúde mental infantil.

Em síntese, este trabalho contribui para o avanço da psiquiatria computacional sob a perspectiva da Engenharia Elétrica, ao demonstrar que a modelagem estatística estruturada de dados psicométricos pode oferecer suporte consistente, interpretável e clinicamente alinhado à decisão em ansiedade pediátrica. As limitações identificadas delineiam caminhos claros para aprofundamento científico, metodológico e translacional, reforçando o potencial do tema para investigações futuras e aplicações clínicas reais.

## Anexo A

### Produção Científica e Registro de Propriedade Intelectual

Os resultados parciais desta dissertação deram origem ao artigo científico de autoria de Renata C. Rocha, Jonathan A. Queiroz e Allan Kardec D. B. Filho, intitulado:

#### *Support System for the Diagnosis of the Risk of Anxiety Disorder in Children*

O referido artigo foi aceito como *artigo estendido* e apresentado oralmente no *XLVI Ibero-Latin American Congress on Computational Methods in Engineering (CILAMCE 2025)*, realizado em Vitória, ES, Brasil, um dos principais congressos internacionais da área de Engenharia Computacional e Métodos Numéricos. A aceitação e a apresentação oral do trabalho atestam a consistência metodológica, a relevância científica e o potencial de aplicação dos resultados desenvolvidos no âmbito desta pesquisa. O artigo encontra-se publicado nos anais do evento, conforme referência a seguir:

ROCHA, Renata C. R. C.; QUEIROZ, Jonathan A.; FILHO, Allan K. D. B. *Support System for the Diagnosis of the Risk of Anxiety Disorder in Children*. In: **XLVI Ibero-Latin American Congress on Computational Methods in Engineering – CILAMCE 2025**, Vitória, ES, Brasil. Anais do CILAMCE 2025. Disponível em: <https://publicacoes.softaliza.com.br/cilamce2025/article/view/13882/8901>.

Adicionalmente, como resultado direto deste trabalho, foi realizado o **Registro de Programa de Computador intitulado “Sistema de Apoio ao Diagnóstico do Risco de Ansiedade Pediátrica”**, junto ao Instituto Nacional da Propriedade Industrial (INPI), sob a linguagem *Python*, consolidando a contribuição tecnológica e a proteção da propriedade intelectual associada ao sistema computacional desenvolvido nesta dissertação.

O registro do programa de computador reforça o caráter aplicado, inovador e translacional da pesquisa, evidenciando seu potencial de utilização em sistemas reais de apoio à decisão clínica em saúde mental pediátrica.

## Anexo B

### Certificado de Aceitação e Apresentação Oral

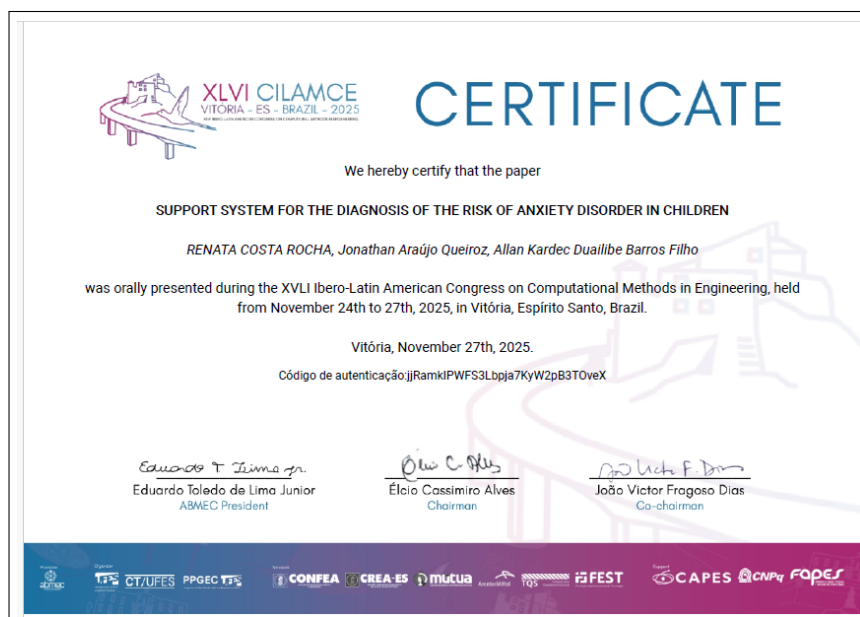
Este anexo apresenta o certificado de participação e apresentação oral do artigo intitulado *Support System for the Diagnosis of the Risk of Anxiety Disorder in Children*, aceito como artigo estendido e apresentado no *CILAMCE 2025 – Ibero-Latin American Congress on Computational Methods in Engineering*.

**Figura 23 – Certificado de aceitação no CILAMCE 2025**



Fonte: Certificado emitido pela organização do CILAMCE 2025.

**Figura 24 – Certificado de apresentação oral no CILAMCE 2025**



Fonte: Certificado emitido pela organização do CILAMCE 2025.

## Anexo C

### Certificado de Registro de Programa de Computador junto ao INPI

Este anexo apresenta o certificado de Registro de Programa de Computador expedido pelo Instituto Nacional da Propriedade Industrial (INPI), referente ao sistema computacional desenvolvido no âmbito desta dissertação, intitulado *Sistema de Apoio ao Diagnóstico do Risco de Ansiedade Pediátrica*. O programa foi registrado sob o número **BR512026000679-1**, cuja concessão do certificado formaliza a proteção da propriedade intelectual do software desenvolvido nesta pesquisa.

**Figura 25 – Certificado de Registro de Programa de Computador expedido pelo INPI**



REPÚBLICA FEDERATIVA DO BRASIL  
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS  
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL  
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

**Certificado de Registro de Programa de Computador**

Processo Nº: **BR512026000679-1**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 26/01/2026, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

Título: SISTEMA DE APOIO AO DIAGNÓSTICO DO RISCO DE ANSIEDADE PEDIÁTRICA

Data de publicação: 26/01/2026

Data de criação: 04/01/2026

Titular(es): Renata Costa Rocha

Autor(es): ALLAN KARDEC DUALIBE BARROS FILHO; JONATHAN ARAUJO QUEIROZ; RENATA COSTA ROCHA

Linguagem: PYTHON

Campo de aplicação: SD-09

Tipo de programa: TC-03

Algoritmo hash: SHA-256

Resumo digital hash: 045AE499D011A2F314FF8A0262277EEBC7E90965BFFF47263393E4F1D9F6CEFD

Expedido em: 10/02/2026

Aprovado por:  
ERICA GUIMARAES CORREA  
Chefe da Divisão de Programas de Computador e Topografias de Circuitos Integrados

Fonte: Instituto Nacional da Propriedade Industrial (INPI), 2026.

## Referências

- 1 VIGO, D. et al. Global burden of mental disorders: The impact of anxiety and depression in children and adolescents. *The Lancet Psychiatry*, v. 7, n. 6, p. 490–499, 2022.
- 2 RACINE, N. et al. Global prevalence of depressive and anxiety symptoms in children and adolescents during covid-19: A meta-analysis. *JAMA Pediatrics*, v. 175, n. 11, p. 1142–1150, 2021.
- 3 HENSE, J. B. et al. Prevalence of common mental disorders in adolescents: a systematic review with meta-analysis. *Jornal de Pediatria*, v. 99, n. 5, p. 453–462, 2023.
- 4 KOPOSOV, R.; MOLEN, M. G. G. T. A. van der; GROOT, P. J. F. de. Developmental trajectories of mental health problems in children and adolescents: a systematic review. *European Child & Adolescent Psychiatry*, v. 30, n. 10, p. 1505–1525, 2021.
- 5 GARCIA-CEJA, E.; OSMANI, V.; MAYORA, O. Machine learning for mental health: A review of current applications and future directions. *Frontiers in Computer Science*, v. 1, p. 1–18, 2018.
- 6 SHATTE, A. B. R.; HUTCHINSON, D. M.; TEAGUE, S. J. Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, Cambridge University Press, v. 49, n. 9, p. 1426–1448, 2019. Disponível em: <https://www.cambridge.org/core/journals/psychological-medicine/article/abs/machine-learning-in-mental-health-a-scoping-review-of-methods-and-applications/0B70B1C827B3A4604C1C01026049F7D9>.
- 7 BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006. ISBN 9780387310732. Disponível em: <https://link.springer.com/book/10.1007/978-0-387-45528-0>.
- 8 HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. *The Elements of Statistical Learning*. [S.l.]: Springer, 2009.
- 9 CARPENTER, K. L. H. et al. Quantifying risk for anxiety disorders in preschool children: A machine learning approach. *PLOS ONE*, Public Library of Science, v. 11, n. 11, p. e0165524, 2016. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0165524>.
- 10 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. 5. ed. Washington, DC: American Psychiatric Publishing, 2013. ISBN 9780890425558.
- 11 World Health Organization. *International Classification of Diseases (ICD-11)*. 11. ed. Geneva: World Health Organization, 2019. [Online]. Disponível em: <https://icd.who.int>. Acesso em: 17 jun. 2025.
- 12 CORTES, C.; VAPNIK, V. Support-vector networks. *Machine Learning*, Springer, v. 20, n. 3, p. 273–297, 1995.
- 13 BREIMAN, L. Random forests. *Machine Learning*, Springer, v. 45, n. 1, p. 5–32, 2001.

- 14 XU, Z. et al. Machine learning approaches for mental health risk prediction: A systematic review. *Artificial Intelligence in Medicine*, v. 141, p. 102558, 2023.
- 15 HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. *Neural Networks*, Elsevier, v. 2, n. 5, p. 359–366, 1989. Disponível em: [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- 16 HAYKIN, S. *Neural Networks and Learning Machines*. [S.l.]: Pearson, 2009.
- 17 LOPES, M. L. *Uso de Redes Neurais Artificiais na identificação de transtornos mentais comuns em crianças e adolescentes*. Dissertação (Dissertação de Mestrado) — Universidade Federal de Pernambuco, 2020.
- 18 KHOEI, T. T.; SLIMANE, H. O.; KAABOUC, N. Deep learning: systematic review, models, challenges, and research directions. *Neural Computing and Applications*, Springer, v. 35, p. 23103–23124, 2023. Disponível em: <https://link.springer.com/article/10.1007/s00521-023-08957-4>.
- 19 STADE, E. C. et al. A transdiagnostic, dimensional classification of anxiety shows improved parsimony and predictive noninferiority to dsm. *Journal of Psychopathology and Clinical Science*, v. 132, n. 8, p. 937–948, 2023. Disponível em: <https://doi.org/10.1037/abn0000863>.
- 20 CASPI, A.; MOFFITT, T. E. All for one and one for all: The p factor and the structure of psychopathology. *Clinical Psychological Science*, v. 8, n. 4, p. 583–595, 2020.
- 21 ARIF, M.; QAMAR, U.; MEHMOOD, R. Machine learning-based identification of generalized anxiety disorder using physiological signals. *IEEE Access*, v. 8, p. 136385–136395, 2020.
- 22 RIVERA, M. J. et al. Diagnosis and prognosis of mental disorders by means of eeg and deep learning: a systematic mapping study. *Artificial Intelligence Review*, Springer, v. 55, p. 1209–1251, 2022. Disponível em: <https://link.springer.com/article/10.1007/s10462-021-09986-y>.
- 23 TALIA, D.; TRUNFIO, P.; PAPUZZO, G. Explainable AI for clinical decision support systems: Challenges and opportunities. *IEEE Transactions on Artificial Intelligence*, v. 5, n. 1, p. 1–14, 2024.
- 24 KAZDIN, A. E.; BLASE, S. L. Rethinking mental health care: Bridging the gap between research and clinical practice. *American Psychologist*, v. 71, n. 7, p. 590–601, 2016.
- 25 CRASKE, M. G. et al. Anxiety disorders. *Nature Reviews Disease Primers*, v. 3, p. 1–18, 2017.
- 26 SMITH, A. e. a. Emotional and socio-cognitive processing in young children with symptoms of anxiety. *Journal of Child Psychology and Psychiatry*, v. 63, n. 10, p. 1123–1134, 2022.
- 27 DICKSON, S. J. et al. Impact of psychotherapy for children and adolescents with anxiety disorders on global and domain-specific functioning: A systematic review and meta-analysis. *Clinical Child and Family Psychology Review*, v. 25, n. 4, p. 720–736, 2022. Disponível em: <https://doi.org/10.1007/s10567-022-00402-7>.
- 28 BARLOW, D. H. *Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic*. New York: Guilford Press, 2002.

- 29 LEDOUX, J. E. Emotion circuits in the brain. *Annual Review of Neuroscience*, v. 23, p. 155–184, 2000.
- 30 DAVIS, M.; WHALEN, P. J. Fear responses and anxiety disorders: neurobiology. *Neuropsychopharmacology*, v. 24, n. 6, p. S13–S34, 2001.
- 31 THAYER, J. F.; LANE, R. D. A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of Affective Disorders*, v. 61, n. 3, p. 201–216, 2000.
- 32 KENDALL, P. C. *Child and Adolescent Therapy: Cognitive-Behavioral Procedures*. New York: Guilford Press, 2002.
- 33 COSTELLO, E. J. et al. Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry*, v. 60, n. 8, p. 837–844, 2003.
- 34 COSTELLO, E. J.; EGGER, H. L.; ANGOLD, A. Epidemiology of anxiety disorders in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America*, v. 14, n. 4, p. 631–648, 2005.
- 35 ANGOLD, A.; COSTELLO, E. J.; ERKANLI, A. Comorbidity. *Journal of Child Psychology and Psychiatry*, v. 43, n. 1, p. 57–87, 2002.
- 36 REYES, A. D. L.; KAZDIN, A. E. Informant discrepancies in the assessment of childhood psychopathology. *Psychological Bulletin*, v. 131, n. 4, p. 483–509, 2005.
- 37 DELCARMEN-WIGGINS, R.; CARTER, A. S. Assessing anxiety disorders in children and adolescents. *Psychiatric Clinics of North America*, v. 28, n. 3, p. 593–616, 2005.
- 38 PINE, D. S.; COHEN, P. Developmental trajectories of anxiety disorders. *Dialogues in Clinical Neuroscience*, v. 11, n. 3, p. 267–276, 2009.
- 39 KRAEMER, H. C. et al. Coming to terms with the terms of risk. *Archives of General Psychiatry*, v. 61, n. 6, p. 625–632, 2004.
- 40 NUNNALLY, J. C.; BERNSTEIN, I. H. *Psychometric Theory*. New York: McGraw-Hill, 2002.
- 41 WEEMS, C. F.; HAYWARD, C.; KILLEN, M. The relation between anxiety sensitivity and internalizing problems in youth. *Journal of Abnormal Child Psychology*, v. 33, n. 1, p. 1–15, 2005.
- 42 ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, American Psychological Association, v. 65, n. 6, p. 386–408, 1958.
- 43 GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. MIT Press, 2016. Disponível em: <<https://www.deeplearningbook.org>>.
- 44 RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, Nature Publishing Group, v. 323, n. 6088, p. 533–536, 1986. Disponível em: <<https://doi.org/10.1038/323533a0>>.
- 45 SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas*. São Paulo: Artliber, 2010.

- 46 STRAUSS, E.; JÚNIOR, M. V. B.; FERREIRA, W. L. L. *A importância de utilizar métricas adequadas de avaliação de performance em modelos preditivos de machine learning*. 2022. Universidade Federal do Rio de Janeiro (UFRJ). Acesso em: jul. 2025. Disponível em: <<https://pdfs.semanticscholar.org/78fa/b9dcbbefa1361fef4a7eb0b849d523aae2aa.pdf>>.
- 47 ZHANG, X. et al. A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *Journal of Neural Engineering*, v. 18, n. 3, p. 031001, 2021. Disponível em: <<https://iopscience.iop.org/article/10.1088/1741-2552/abdfe0>>.
- 48 KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, 2015. Disponível em: <<https://arxiv.org/abs/1412.6980>>.
- 49 SCHÖLKOPF, B.; SMOLA, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press, 2002.
- 50 IA com Café. *Implementando SVM do zero*. 2023. Acesso em: 4 jan. 2026. Disponível em: <<https://iacomcafe.com.br/implementando-svm-do-zero/>>.
- 51 Statplace. *Modelos de Machine Learning utilizando o pacote caret*. 2020. Acesso em: 4 jan. 2026. Disponível em: <<https://statplace.com.br/blog/modelos-de-machine-learning-utilizando-o-pacote-caret/>>.
- 52 DOSHI-VELEZ, F.; KIM, B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017. Disponível em: <<https://arxiv.org/abs/1702.08608>>.
- 53 LIPTON, Z. C. The mythos of model interpretability. *arXiv preprint arXiv:1606.03490*, 2018. Disponível em: <<https://arxiv.org/abs/1606.03490>>.
- 54 LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, Curran Associates, Inc., v. 30, p. 4765–4774, 2017. Disponível em: <[https://proceedings.neurips.cc/paper\\_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html](https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html)>.
- 55 POWERS, D. M. W. Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, v. 2, n. 1, p. 37–63, 2011. Disponível em: <<https://www.bioinf.jku.at/publications/older/2604.pdf>>.
- 56 RASCHKA, S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, 2018. Disponível em: <<https://arxiv.org/abs/1811.12808>>.
- 57 NOGARE, D. *Performance de Machine Learning: Matriz de Confusão*. 2020. Disponível em: <<https://diegonogare.net/2020/04/performance-de-machine-learning-matriz-de-confusao/>>.
- 58 JOHNSON, A.; KHOSHGOFTAAR, T. M. Survey on deep learning with class imbalance. *Journal of Big Data*, v. 9, n. 1, p. 1–56, 2022. Disponível em: <<https://doi.org/10.1186/s40537-022-00608-7>>.
- 59 NAY, N. *Mastering Evaluation Metrics: F1 Score, Z-Score, ROC Curve and Precision-Recall Curve*. 2020. Disponível em: <<https://medium.com/@nay1228/mastering-evaluation-metrics-f1-score-z-score-roc-curve-and-precision-recall-curve-cb175de25724>>.

- 60 FAWCETT, T. An introduction to roc analysis. *Pattern Recognition Letters*, v. 27, n. 8, p. 861–874, 2006.
- 61 Evidently AI. *ROC Curve Explained*. 2023. Disponível em: <<https://www.evidentlyai.com/classification-metrics/explain-roc-curve>>.
- 62 JAMES, G. et al. *An Introduction to Statistical Learning with Applications in Python*. New York: Springer, 2021.
- 63 RASCHKA, S.; LIU, Y.; MIRJALILI, V. *Machine Learning with PyTorch and Scikit-Learn*. Birmingham, UK: Packt Publishing, 2020.
- 64 Scikit-learn Developers. *Cross-validation: evaluating estimator performance*. 2023. <[https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html)>. Acesso em: 2026.
- 65 KUHN, M.; SILGE, J. *Tidy Modeling with R*. Sebastopol: O’Reilly Media, 2022.
- 66 IA com Café. *A importância da validação cruzada em Machine Learning*. 2023. <<https://iacomcafe.com.br/importancia-validacao-cruzada-machine-learning/>>. Acesso em: 2026.
- 67 CARPENTER, K. 2016. Disponível em: <<https://doi.org/10.7910/DVN/N42LWG>>.
- 68 WILKINSON, M. D. et al. The fair guiding principles for scientific data management and stewardship. *Scientific Data*, Nature Publishing Group, v. 3, n. 1, p. 1–9, 2016. Disponível em: <<https://doi.org/10.1038/sdata.2016.18>>.
- 69 ALLEN, J. L.; EHRENREICH-MAY, J.; BARLOW, D. H. The development of anxiety disorders: Phenomenology, classification, and developmental pathways. *Journal of Child Psychology and Psychiatry*, v. 60, n. 4, p. 333–349, 2019.
- 70 STEYERBERG, E. W. *Clinical Prediction Models*. [S.l.]: Springer, 2019.
- 71 GRINSZTAJN, F.; OYALLON, E.; VAROQUAUX, G. Why do tree-based models still outperform deep learning on tabular data? *arXiv preprint arXiv:2207.08815*, 2022.
- 72 TOPOL, E. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. [S.l.]: Basic Books, 2019.
- 73 HOLZINGER, A. et al. Explainable ai systems in healthcare. *Artificial Intelligence in Medicine*, 2019.
- 74 World Health Organization and UNICEF. *Guidance to improve access to mental health care for children and young people*. 2024. Disponível em: <<https://www.who.int/news/item/09-10-2024-who-and-unicef-launch-guidance-to-improve-access-to-mental-health-care-for-children-and-yo>>
- 75 WANG, S. et al. Treatment rates for mental disorders among children and adolescents: A systematic review and meta-analysis. *JAMA Network Open*, v. 6, n. 10, p. e2338174, 2023.
- 76 BONE, D. et al. Ethical and methodological concerns in ai for mental health. *Journal of Child Psychology and Psychiatry*, 2017.