



Modelo preditivo para resultados clínicos da tuberculose com Redes Neurais

Predictive model for tuberculosis clinical outcomes with neural networks

Modelo predictivo para resultados clínicos de tuberculosis con redes neuronales

Ronilson Williame da Silva Pereira¹, Igor Wenner Silva Falcão², Saul Carneiro³,
Marcos César da Rocha Seruffo², Karla Figueiredo⁴

RESUMO

Descriptores: tuberculose; modelagem preditiva com aprendizado de máquina; MLP.

Objetivo: Este estudo propõe o desenvolvimento de um modelo baseado em Redes Neurais do tipo Multilayer Perceptron (MLP) para prever desfechos no tratamento da tuberculose, com ênfase na identificação de casos de cura e abandono. **Métodos:** Os dados foram pré-processados com imputação de ausentes, codificação categórica e normalização. Utilizou-se a técnica SMOTETomek para balanceamento. A arquitetura da MLP incluiu camadas densas com ativação ReLU, regularização com dropout (50%) e saída sigmoide. O modelo foi treinado com validação e avaliado com e sem balanceamento. **Resultados:** No cenário sem balanceamento, o modelo obteve acurácia macro de 0,6235, precisão de 0,9115, recall de 0,9781 e F1-macro de 0,6584, indicando viés em favor da majoritária. Com balanceamento, a acurácia micro e o F1-micro atingiram 0,8571. A precisão foi de 0,8857 e o recall reduziu-se para 0,8197. Na análise por classe, o modelo apresentou melhor desempenho ao prever abandono (F1 = 0,8623) em comparação com cura (F1 = 0,8514). **Conclusões:** O balanceamento de classes contribuiu para a melhoria do desempenho geral do modelo. O uso de MLP, aliado a estratégias técnicas de pré-processamento e balanceamento, mostrou-se eficaz na predição de desfechos no tratamento da tuberculose.

ABSTRACT

Descriptors: tuberculosis; predictive modeling with machine learning; MLP.

Objective: This study proposes the development of a model based on Multilayer Perceptron (MLP) neural networks to predict outcomes in tuberculosis treatment, focusing on identifying cases of cure and abandonment. **Methods:** Data were preprocessed using missing value imputation, categorical encoding, and normalization. The SMOTETomek technique was applied for balancing. The MLP architecture included dense layers with ReLU activation, 50% dropout regularization, and sigmoid output. The model was trained with validation and evaluated in both balanced and unbalanced scenarios. **Results:** In the unbalanced scenario, the model achieved macro accuracy of 0.6235, precision of 0.9115, recall of 0.9781, and F1-macro of 0.6584, indicating bias toward the majority class. With balancing, micro accuracy and F1-micro reached 0.8571. Precision was 0.8857, while recall dropped to 0.8197. In class-specific analysis, the model performed better for abandonment (F1 = 0.8623) compared to cure (F1 = 0.8514). **Conclusion:** Class balancing improved the model's overall performance. The application of MLP, combined with preprocessing and balancing strategies, proved effective for predicting outcomes in tuberculosis treatment.

RESUMEN

Descriptores: tuberculosis; modelado predictivo con aprendizaje automático; MLP.

Objetivo: Este estudio propone el desarrollo de un modelo basado en redes neuronales del tipo Multilayer Perceptron (MLP) para predecir desenlaces en el tratamiento de la tuberculosis, con énfasis en la identificación de casos de curación y abandono. **Métodos:** Los datos fueron preprocesados mediante imputación de valores faltantes, codificación categórica y normalización. Se aplicó la técnica SMOTETomek para el balanceo. La arquitectura de la MLP incluyó capas densas con activación ReLU, regularización con dropout (50%) y salida sigmoide. El modelo fue entrenado con validación y evaluado en escenarios con y sin balanceo. **Resultados:** En el escenario sin balanceo, el modelo obtuvo una precisión macro de 0,6235, precisión de 0,9115, recall de 0,9781 y F1-macro de 0,6584, indicando sesgo hacia la clase mayoritaria. Con balanceo, la precisión micro y el F1-micro alcanzaron 0,8571. La precisión fue de 0,8857 y el recall bajó a 0,8197. En el análisis por clase, el modelo mostró mejor rendimiento al predecir abandono (F1 = 0,8623) en comparación con curación (F1 = 0,8514). **Conclusión:** El balanceo de clases mejoró el rendimiento general del modelo. El uso de MLP, junto con estrategias de preprocesamiento y balanceo, fue eficaz para predecir desenlaces en el tratamiento de la tuberculosis.

¹Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

²Instituto de Tecnologia, Universidade Federal do Pará, Belém, PA, Brasil.

³Hospital Universitário João de Barros Barreto, Universidade Federal do Pará, Belém, PA, Brasil.

⁴Departamento de Ciência da Computação, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

INTRODUÇÃO

A tuberculose (TB) é uma doença infecciosa causada pela bactéria *Mycobacterium tuberculosis*, que afeta principalmente os pulmões humanos, e é considerada um dos principais desafios de saúde pública global⁽¹⁾. De acordo com o Relatório Global sobre Tuberculose de 2024 da Organização Mundial da Saúde (OMS), milhares de pessoas são diagnosticadas com TB anualmente, e uma parcela significativa dos casos evolui para o abandono do tratamento, comprometendo a eficácia das intervenções médicas e contribuindo para a disseminação da doença⁽²⁾.

O sucesso do tratamento da TB depende de diversos fatores, como adesão à medicação, resistência bacteriana, presença de comorbidades (por exemplo: HIV, diabetes), condições socioeconômicas e características clínicas dos pacientes⁽³⁾. A complexidade do regime terapêutico que geralmente envolve múltiplas drogas e um longo período de administração torna a predição de desfechos clínicos uma tarefa desafiadora, porém fundamental para orientar intervenções precoces.

Nesse contexto, técnicas de Inteligência Artificial (IA), em especial do campo de Aprendizado de Máquina (*Machine Learning* - ML), têm ganhado destaque em aplicações na área da saúde. As Redes Neurais Artificiais (RNAs), como o *Perceptron Multicamadas* (*Multilayer Perceptron* - MLP), são capazes de modelar relações complexas entre variáveis e tem sido amplamente utilizada em tarefas de classificação binária com bons resultados⁽⁴⁾. O MLP é composto por múltiplas camadas ocultas com funções de ativação não lineares, permitindo alta capacidade de generalização e aprendizado de padrões não lineares nos dados⁽⁵⁾.

A escolha do MLP neste estudo fundamenta-se em sua eficácia demonstrada na literatura em contextos similares, superando, frequentemente, modelos clássicos como regressão logística e *Naive Bayes*⁽⁶⁾. Além disso, a arquitetura MLP oferece flexibilidade para ajustes conforme a natureza dos dados clínicos e permite integração eficiente com técnicas de balanceamento de classes⁽⁷⁾, como o *SMOTETomek*.

Este trabalho propõe o desenvolvimento de um modelo preditivo baseado em MLP para prever os desfechos do tratamento da tuberculose, distinguindo entre cura e abandono. Foram aplicadas técnicas de pré-processamento, regularização por *dropout* e balanceamento de classes com *SMOTETomek*, visando melhorar o desempenho do modelo e reduzir o viés em relação à classe majoritária.

Embora as redes MLP sejam uma técnica bem estabelecida, nossa abordagem se destaca ao combinar MLP e *SMOTETomek* para o problema dos desfechos do tratamento da tuberculose, oferecendo uma solução específica para lidar com o desequilíbrio de dados.

Diferentemente das abordagens tradicionais, nosso estudo não apenas avalia o desempenho do MLP, mas também investiga o impacto do balanceamento de dados na precisão e sensibilidade do modelo, especialmente em dados médicos, onde é crucial identificar desistências.

O restante do artigo é organizado da seguinte forma: os principais trabalhos correlatos são detalhados na Seção 2. A Seção 3 apresenta a metodologia utilizada, detalhando o modelo proposto. A seção 4 apresenta os resultados obtidos para a previsão de desfecho de tratamento. Por fim, as principais conclusões alcançadas são resumidas e os trabalhos futuros explicados na Seção 5.

TRABALHOS CORRELATOS

Diversos estudos têm explorado o uso de técnicas de aprendizado de máquina na tuberculose, tanto para diagnóstico quanto para a predição de desfechos. Esta seção discute os principais trabalhos relacionados e destaca como a presente proposta se diferencia.

No estudo de Liao et al. (2023)⁽⁴⁾, os autores propuseram modelos preditivos para hepatite aguda, insuficiência respiratória aguda e mortalidade após o tratamento da tuberculose, utilizando seis algoritmos de aprendizado de máquina (*XGBoost*, *random forest*, MLP, *light GBM*, regressão logística e SVM). Os resultados indicaram que o MLP obteve o maior valor de área sob a curva (AUC) (0,834) para a predição de mortalidade.

No trabalho de Janah et al. (2024)⁽⁸⁾, os autores desenvolveram modelos MLP e *Extreme Learning Machine* (ELM) para detecção precoce de TB com dados clínicos. Embora tenham aplicado a técnica de sobreamostragem da classe minoritária (*Synthetic Minority Oversampling Technique* - SMOTE) para balanceamento, o melhor resultado (acurácia de 95%) foi obtido com MLP sem a técnica de balanceamento, indicando que a eficácia do SMOTE pode variar conforme o problema e os dados utilizados.

No estudo de Mohidem et al. (2021)⁽⁹⁾, os autores desenvolveram um modelo para prever o número de casos de TB no distrito de Gombak, na Malásia, com base em fatores sociodemográficos e ambientais. Foram utilizadas Regressão Linear Múltipla e Redes Neurais Artificiais (RNA) para desenvolver o modelo. A abordagem com melhor desempenho foi a RNA3, que avaliou a relação entre os aspectos sociodemográficos e os fatores ambientais, alcançando uma acurácia de 96%.

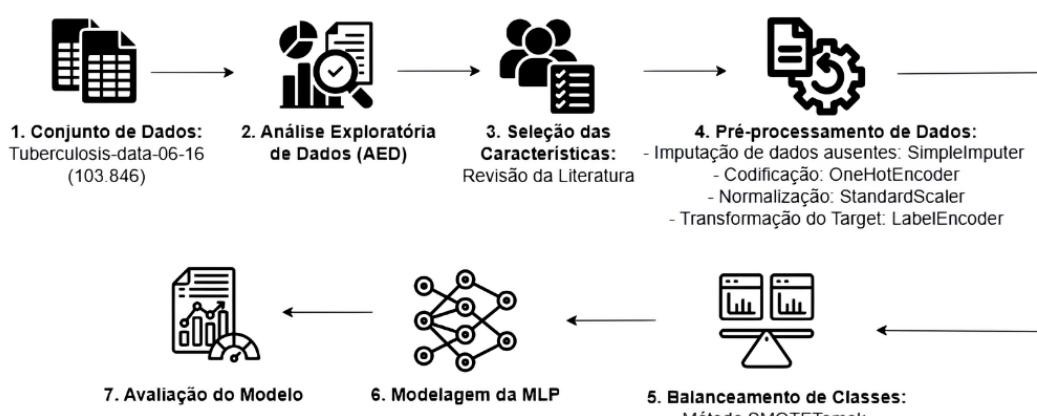
Um estudo de Orjuela et al. (2022)⁽⁶⁾ explorou o uso de aprendizado de máquina para o diagnóstico de tuberculose em cenários com recursos limitados. Foram desenvolvidos modelos com técnicas como regressão logística, árvores de classificação, *random forest*, SVM e redes neurais artificiais. O modelo de RNA obteve os melhores resultados, apresentando melhores resultados em termos de acurácia, sensibilidade e AUC.

Este trabalho propõe a aplicação da arquitetura MLP para prever desfechos clínicos binários em casos de tuberculose (cura vs. abandono), integrando a técnica *SMOTETomek* para o balanceamento das classes. O modelo foi desenvolvido com base em 103.846 registros reais do sistema TBWEB, incluindo variáveis clínicas e socioeconômicas. O estudo também avalia o impacto do balanceamento nas métricas de acurácia, precisão e *recall*, com ênfase no compromisso entre essas métricas, especialmente em contextos críticos de saúde pública.

MATERIAS E MÉTODOS

Nesta seção, são detalhadas as etapas e os métodos utilizados para o desenvolvimento do modelo para a previsão dos desfechos de tratamento da tuberculose. Detalharam-se os passos a definição do conjunto de dados, passando pelo pré-processamento, modelagem e avaliação dos dados, visando ao desenvolvimento de um modelo robusto. O fluxo de trabalho metodológico utilizado neste estudo está resumido na Figura 1.

Figura 1. Fluxo de trabalho da metodologia utilizada.



Fonte: Elaborado pelos próprios autores.

Conjunto de Dados

A base de dados utilizada está disponível em repositório público. Este estudo analisou dados de 103.846 prontuários de pacientes com tuberculose no estado de São Paulo, obtidos entre 2006 a 2016 por meio do TBWEB, sistema de informação utilizado para monitoramento da tuberculose. O conjunto de dados inclui variáveis demográficas, clínicas e de tratamento.

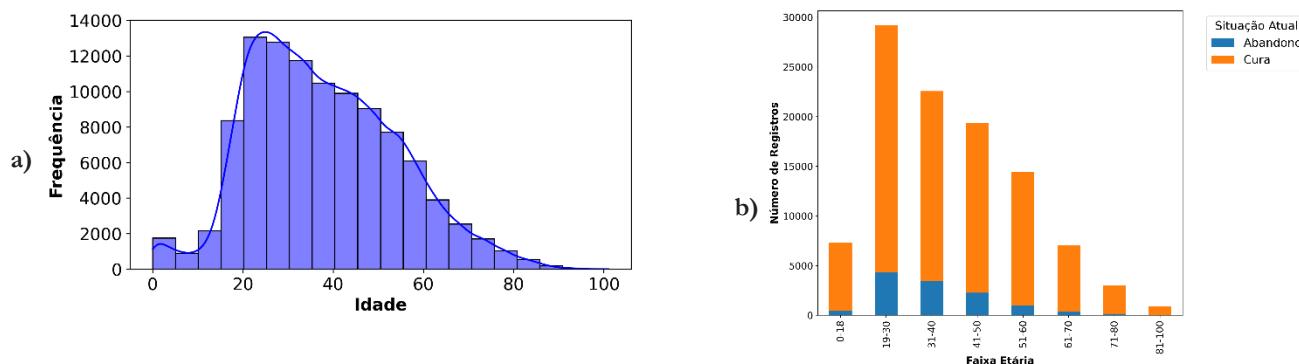
Análise Exploratória de Dados (AED)

A Análise Exploratória de Dados (AED) tem como objetivo examinar o conjunto de dados antes da aplicação de técnicas de aprendizado de máquina ou métodos estatísticos. Essa etapa é essencial para compreender a distribuição dos dados e as características dos atributos disponíveis, identificar possíveis anomalias e orientar decisões no pré-processamento.

Foram geradas representações gráficas apropriadas para os diferentes tipos de variáveis: histogramas para variáveis numéricas e gráficos de barras para variáveis categóricas. Entre as variáveis analisadas, destacam-se a distribuição etária dos pacientes, a distribuição por sexo e a situação atual do tratamento (cura ou abandono).

A Figura 2 apresenta duas visualizações: (a) a distribuição etária dos pacientes no conjunto de dados e (b) o número de registros por faixa etária, discriminados pela situação atual do tratamento. A análise mostra que a maior concentração de casos está na faixa de 20 a 25 anos. Há uma redução progressiva na frequência de registros a partir dos 60 anos, e observa-se uma baixa incidência de casos em crianças entre 5 e 10 anos.

Na análise cruzada entre faixa etária e desfecho do tratamento, observa-se que o maior número de pacientes curados encontra-se na faixa de 19 a 30 anos, com destaque para o total de 8.782 pacientes curados nesse grupo. Esse comportamento reflete, em parte, a maior incidência de TB em adultos jovens. Os dados também revelam um desbalanceamento acentuado entre as classes de desfecho: a classe “cura” predomina amplamente sobre a classe “abandono”, o que reforça a necessidade da aplicação de técnicas de balanceamento durante o processo de modelagem.

Figura 2. Distribuição etária dos pacientes (a); número de registros por faixa etária segundo a situação atual (b).

Fonte: Elaborado pelos próprios autores.

Seleção das Características

Para facilitar a análise, as variáveis preditoras foram categorizadas com base em características (*features*) sociais e clínicas. A análise foi fundamentada na revisão da literatura, na disponibilidade de dados e na importância para a previsão dos desfechos, conforme apresentado nos trabalhos de Orjuela et al. (2022)⁽⁶⁾ e Kanesamoorthy et al. (2021)⁽¹⁰⁾.

A Tabela 1 apresenta os atributos clínicos utilizados no modelo, incluindo comorbidades relevantes como diabetes, HIV, alcoolismo e dependência química, além de exames diagnósticos como (bacilosкопия, raio-X, histopatologia) e variáveis associadas à condução do tratamento, como tipo

de instituição, esquema inicial e forma clínica. Indicadores como o número de doses recebidas refletem a adesão ao tratamento, enquanto variáveis como idade e tipo de caso são essenciais para o planejamento terapêutico.

O conjunto final de dados compreendeu 24 atributos, agrupados em duas categorias: clínicos (Tabela 1) e socioeconômicos (Tabela 2). Foram selecionadas quatro características socioeconômicas: raça/cor, sexo, escolaridade e tipo de ocupação, conforme ilustrado na Tabela 2. Como apontado por Rodrigues et al. (2020)⁽¹¹⁾, fatores demográficos e sociais influenciam significativamente a adesão ao tratamento e os desfechos clínicos em pacientes com tuberculose.

Tabela 1. Atributos clínicos selecionados (Agrupados por categorias).

Categoria	Atributo	Descrição	Valores possíveis
Diagnóstico	tipoCaso	Tipo de caso de TB	Novo, Recidiva, Reingresso após abandono, Reingresso após falha
	FORMACLIN1	Forma clínica da TB	Pulmonar, Extrapulmonar, Miliar, Ganglionar
	classif	Classificação do caso	Confirmado, Provável, Descartado
	Bac	Baciloscopy de escarro	Negativo, Positivo, Não realizado
Exames	BACOUTRO	Baciloscopy de outro material	Negativo, Positivo, Não realizado
	cultEsc	Cultura de escarro	Negativo, Positivo, Em andamento
	RX	Radiografia do tórax	Normal, Suspeito de TB, Outra doença
	HISTOPATOL	Exame histopatológico	Compatível com TB, Inconclusivo
Comorbidades	NECROP	Resultado de necropsia	BAAR positivo, Suspeito de TB, Não realizado
	hiv	Status HIV	Positivo, Negativo, Não testado
	DIABETES	Diagnóstico de diabetes	Sim, Não
	ALCOOLISMO	Uso problemático de álcool	Sim, Não
Tratamento	MENTAL	Doença mental diagnosticada	Sim, Não
	DROGADICAO	Uso de drogas ilícitas	Sim, Não
	TABAGISMO	Hábito de fumar	Sim, Não
	instTrat	Tipo de instituição de tratamento	Ambulatorial, Hospitalar, Prisional
Demográficos	esqIni	Esquema terapêutico inicial	RHZE, RHZ, Outros
	nDosesPri	Número de doses (1º-2º mês)	Valor numérico
Desfecho	sitAtual	Situação atual do tratamento	Abandono, Cura

Fonte: Elaborado pelos próprios autores.

Tabela 2. Atributos socioeconômicos selecionados.

Atributos Sociais	Descrição	Valores possíveis
racaCor	Auto identificação étnica	Branca, Preta, Parda, Amarela, Indígena
sexo	Sexo biológico	Masculino, Feminino
ESCOLARID	Escolaridade máxima completada	Nenhuma, 1-3 anos, 4-7 anos, 8-11 anos, ≥12 anos
TIPOCUP	Ocupação principal	Desempregado, Aposentado, Profissional de saúde, Outros

Fonte: Elaborado pelos próprios autores.

Pré-processamento dos Dados

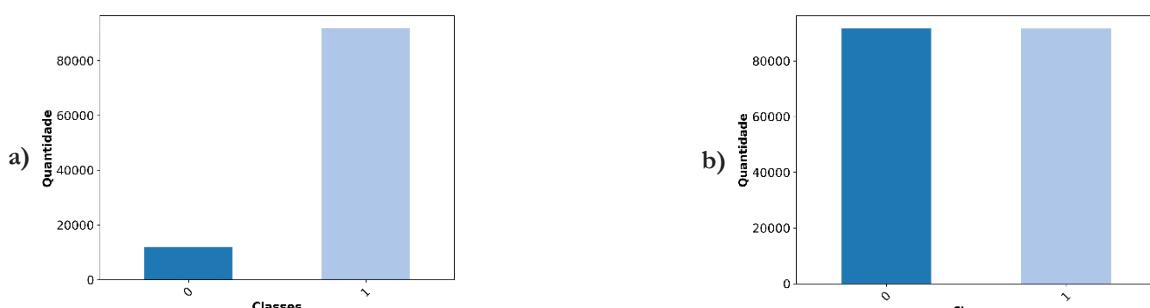
Durante a etapa de pré-processamento, identificou-se a presença de valores ausentes em variáveis como escolaridade, tipo de ocupação, forma clínica, comorbidades (por exemplo: diabetes, alcoolismo) e tipo de instituição. As variáveis categóricas foram imputadas com o valor mais frequente (moda), enquanto as numéricas (como idade) foram preenchidas com a mediana, seguindo abordagem padrão em problemas de saúde pública⁽¹²⁾.

A codificação de variáveis categóricas foi realizada por meio do método *One-Hot Encoding*, transformando cada categoria em uma coluna binária. Já as variáveis numéricas foram normalizadas com o *StandardScaler*, que ajusta os valores para média zero e desvio padrão um. A variável alvo foi convertida em binária: ‘0’ para abandono e ‘1’ para cura, totalizando 91.823 registros como desfecho de cura e 12.023 de abandono.

Balanceamento de Classes

Devido ao desbalanceamento severo da variável alvo (proporção de aproximadamente 7,6:1 entre as classes), aplicou-se a técnica *SMOTETomek*, que combina a geração sintética de amostras minoritárias (*SMOTE*) com a remoção de pares ambíguos (*Tomek Links*), conforme proposto por *Talukder et al.* (2024)⁽¹³⁾. Conforme verificado na fase de AED e ilustrado na Figura 3 a), utilizou-se a técnica *SMOTETomek* para resolver esse desequilíbrio. Essa abordagem promove um balanceamento mais refinado e contribui para reduzir o viés do modelo na classe majoritária. A Figura 3 (b) mostra o resultado do balanceamento, aumentando o número de registros da classe ‘0’ e promovendo uma distribuição mais equilibrada entre as classes ‘0’ e ‘1’.

Figura 3. (a) Distribuição da classe de saída antes do balanceamento dos dados; (b) Distribuição da classe de saída após o balanceamento dos dados com SMOTETomek.



Fonte: Elaborado pelos próprios autores.

Modelagem e Avaliação

Nesta seção, apresenta-se o modelo de aprendizado de máquina utilizado para prever os desfechos clínicos em pacientes com tuberculose: o Perceptron Multicamadas (Multilayer Perceptron - MLP). A escolha pelo MLP justifica-se por sua robustez, capacidade de generalização e bom desempenho em tarefas com dados estruturados, especialmente em contextos clínicos com variáveis complexas e interdependentes.

A modelagem foi conduzida na linguagem Python (versão 3.13.2), utilizando a biblioteca tensorflow.keras. O fluxo de desenvolvimento seguiu um pipeline estruturado, com divisão estratificada dos dados平衡ados: 80% para treino e 20% para teste. Durante o treinamento, 20% do conjunto de treino foi destinado à validação.

A arquitetura final da rede foi definida com base em experimentos preliminares e recomendações da literatura⁽¹⁴⁾, utilizando uma abordagem de ajuste manual de hiperparâmetros (manual tuning). O modelo MLP consistiu em três camadas densas intercaladas por camadas de regularização via dropout (taxa de 50%). As camadas oculadas possuíam 128 e 64 neurônios, respectivamente, ambas com função de ativação ReLU (Rectified Linear Unit). A camada de saída, com um único neurônio e ativação sigmoide, foi utilizada para realizar a predição binária entre cura e abandono do tratamento.

A MLP implementada é composta por camadas densas intercaladas com regularização via dropout e uma camada de saída com ativação sigmoide. A rede totalizou 22.913 parâmetros treináveis, sem a presença de parâmetros não

ajustáveis, indicando que toda a arquitetura contribui efetivamente para o processo de aprendizado. A ausência de camadas convolucionais ou mecanismos especializados mantém o modelo simples e eficiente, favorecendo sua aplicação em contextos clínicos com recursos computacionais limitados. O fluxo unidirecional entre as camadas (feedforward), característico das redes MLP, é preservado em toda a estrutura, garantindo a propagação direta das informações durante o treinamento e a inferência.

A Tabela 3 apresenta os principais hiperparâmetros utilizados no modelo, incluindo número de camadas, taxa de dropout, funções de ativação, otimizador, número de épocas e tamanho de lote (batch size). Esses parâmetros foram definidos com base no desempenho observado nos dados de validação, priorizando o equilíbrio entre acurácia e capacidade de generalização.

Tabela 3. Hiperparâmetros do modelo MLP

Modelo	Hiperparâmetros
	- Camadas ocultas: [128, 64]
	- <i>Dropout</i> : 0,5
MLP	- Função de ativação: ReLU
	- Otimizador: Adam (<i>learning rate</i> = 0,001)
	- Épocas: 50
	- Tamanho de lote (<i>batch size</i>): 32

Fonte: Elaborado pelos próprios autores.

A avaliação do modelo considerou dois cenários experimentais: (i) treinamento com o conjunto original,

sem balanceamento das classes; e (ii) treinamento com o conjunto balanceado utilizando a técnica SMOTETomek. Para cada condição, foram aplicadas as métricas de acurácia, precisão, recall e F1-score.

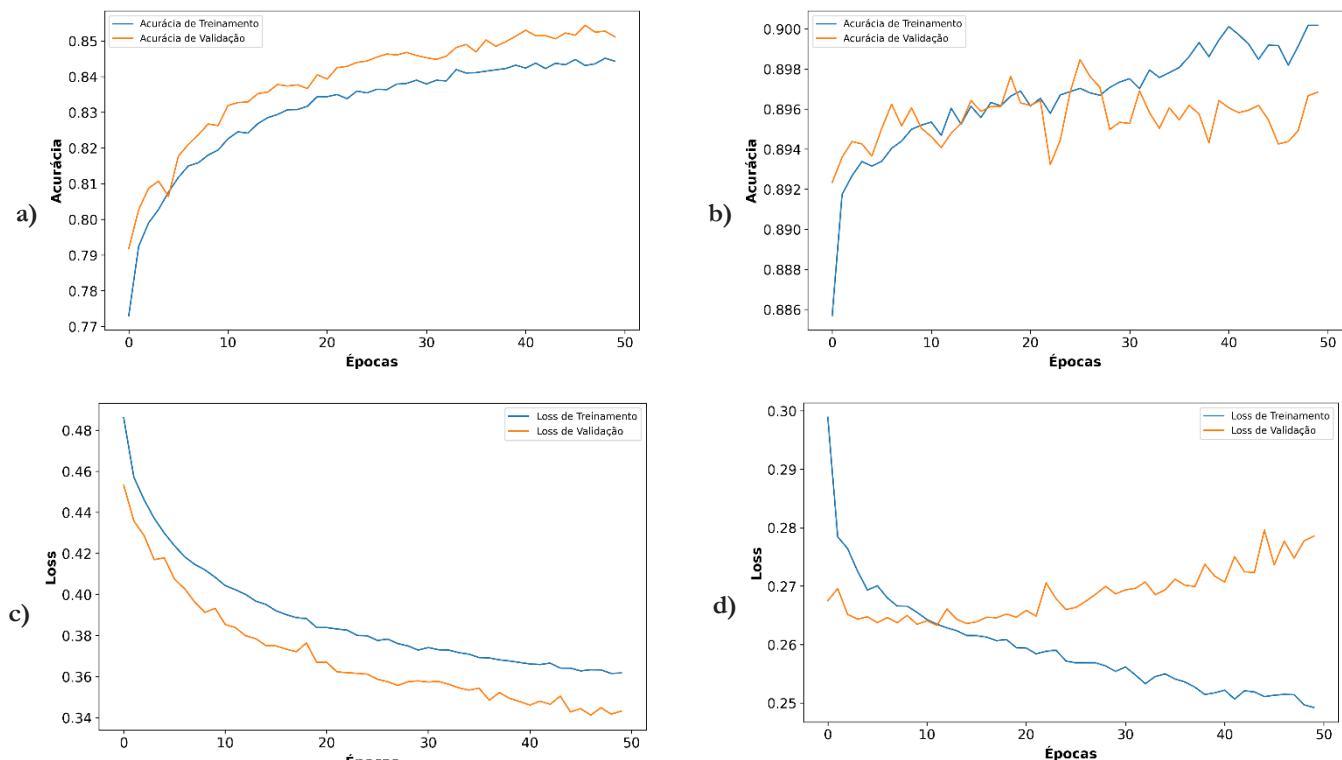
Nos dados desbalanceados, priorizaram-se as métricas macro, como F1-macro e acurácia macro, que atribuem peso igual a cada classe e ajudam a mitigar o viés a favor da classe majoritária. Já nos dados balanceados, foram utilizadas as métricas micro (F1-micro e acurácia micro), apropriadas para classes com número de registros semelhantes, pois refletem o desempenho geral do modelo. Além dessas métricas, foram analisadas as curvas de aprendizado (perda e acurácia por época) e as matrizes de confusão, permitindo avaliar com maior precisão os erros cometidos em cada classe.

RESULTADOS E DISCUSSÃO

Esta seção apresenta os resultados obtidos pelo modelo MLP, nas versões com e sem balanceamento de classes, discutindo o impacto do desbalanceamento sobre a performance e a capacidade de generalização.

A Figura 4 exibe as curvas de acurácia e perda durante as épocas de treinamento. Com os dados平衡ados (Figura 6d), observa-se uma tendência consistente de redução da perda tanto em treino quanto na validação. No modelo sem balanceamento (Figura 6c), a perda de validação apresenta maior variabilidade, sugerindo um possível overfitting.

Figura 4. (a) Acurácia ao longo das épocas sem balanceamento; (b) Acurácia ao longo das épocas com balanceamento; (c) Curva de aprendizado sem balanceamento; (d) Curva de aprendizado com balanceamento.



Fonte: Elaborado pelos próprios autores.

A Tabela 4 apresenta os principais resultados agregados. No cenário sem balanceamento, o modelo obteve acurácia macro de 0,6235, com precisão de 0,9115 e recall de 0,9781. Entretanto, como esperado, o desempenho foi influenciado

pela predominância da classe majoritária (cura). Após o balanceamento com SMOTETomek, a acurácia micro aumentou para 0,8571, com precisão de 0,8857 e recall de 0,8197 e F1-micro de 0,8571, indicando melhor equilíbrio geral.

Tabela 4. Resultados das Métricas de Avaliação.

Métrica	Conjunto Não Balanceado	Conjunto Balanceado (<i>SMOTETomek</i>)
Acurácia Macro	0,6235	-
Precisão	0,9115	0,8857
Recall	0,9781	0,8197
F1-macro	0,6584	-
Acurácia micro	-	0,8571
F1-micro	-	0,8571

Fonte: Elaborado pelos próprios autores.

A Tabela 5 apresenta as métricas desagregadas por classe (cura e abandono), permitindo avaliar o desempenho do modelo especificamente na classe minoritária, de maior interesse em saúde pública.

Tabela 5. Métricas por Classe (modelo com balanceamento).

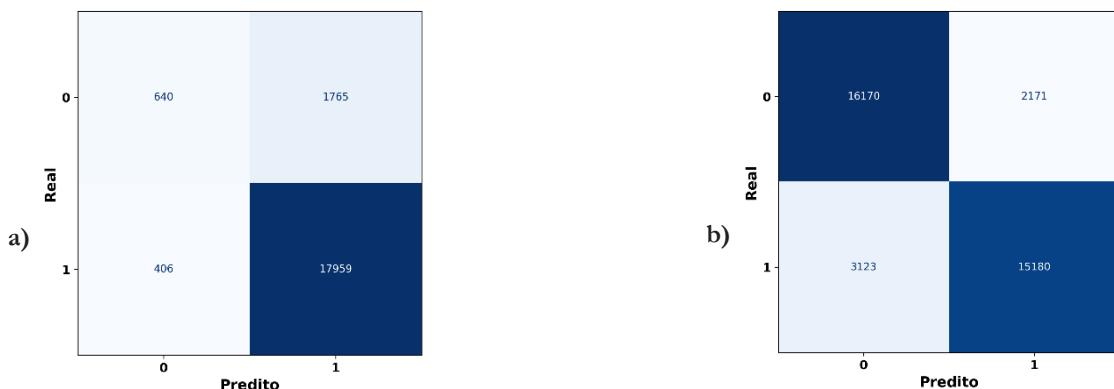
Classe	Precisão	Recall	F1-score
Abandono (0)	0,8325	0,8944	0,8623
Cura (1)	0,8857	0,8197	0,8514

Fonte: Elaborado pelos próprios autores.

Com o balanceamento, observa-se uma leve redução no *recall* para a classe abandono (0,75), o que indica que o modelo ainda possui limitações na identificação de pacientes com maior risco de não adesão. Esse comprometimento entre precisão e sensibilidade é comum em problemas clínicos e deve ser cuidadosamente considerado

na aplicação prática dos modelos. A Figura 5 apresenta as matrizes de confusão para os cenários com e sem balanceamento. No conjunto não balanceado, o modelo apresenta desempenho elevado para a classe cura, mas falha em identificar abandonos. Após o balanceamento, a distribuição dos acertos entre as classes torna-se mais equilibrada.

Figura 5. a) Matriz de Confusão sem Balanceamento; (b) Matriz de Confusão com Balanceamento.



Fonte: Elaborado pelos próprios autores.

CONCLUSÃO

Este estudo propôs o desenvolvimento e avaliação de um modelo preditivo baseado em redes neurais *Multilayer Perceptron* (MLP) para prever desfechos no tratamento da tuberculose, com foco na diferenciação entre cura e abandono. A abordagem incluiu pré-processamento de dados clínicos e socioeconômicos, balanceamento de classes com *SMOTETomek* e regularização por *dropout*, utilizando dados reais do sistema TBWEB, com 103.846 prontuários

de pacientes com tuberculose no estado de São Paulo.

Os resultados indicaram que o modelo MLP obteve desempenho satisfatório na tarefa de classificação binária, com acurácia *micro* de 0,8571 após o balanceamento de classes, e F1-score de 0,8571. A aplicação da técnica *SMOTETomek* contribuiu para reduzir o viés em favor da classe majoritária, permitindo maior sensibilidade na identificação de pacientes com risco de abandono, ainda que com um comprometimento na precisão.

Embora o uso de atributos clínicos e socioeconômicos tenha possibilitado o treinamento de um modelo com bom desempenho, não foi realizada neste estudo uma análise formal da importância das variáveis. Assim, quaisquer afirmações sobre as características que mais impactaram a predição devem ser interpretadas com cautela. Essa limitação deve ser considerada ao interpretar os resultados do estudo.

Como trabalhos futuros, propõe-se aplicar métodos de explicabilidade, como SHAP e análise de importância por permutação, para identificar variáveis-chave na decisão do modelo, comparar o desempenho do MLP com modelos *baseline*, como regressão logística, árvore de decisão e SVM, para justificar sua superioridade, avaliar arquiteturas alternativas (como redes profundas ou híbridas) e explorar validação cruzada com *tuning* automático de hiperparâmetros, expandir a análise para bases nacionais (ex.: SINAN), a fim de verificar a generalização do modelo em outras populações.

REFERÊNCIAS

1. Obeagu E, Obeagu G. Understanding immune cell trafficking in tuberculosis-hiv coinfection: The role of l-selectin pathways. *Elite Journal of Immunology*. 2024; 2(2):43–59.
2. WHO (2024). Global Tuberculosis Report. Number September.
3. Motta I, Boeree M, Chesov D, Dheda K, Günther G, Horsburgh Jr CR, et al. Recent advances in the treatment of tuberculosis. *Clinical Microbiology and Infection*. 2023.
4. Liao KM, Liu CF, Chen CJ, Feng JY, Shu CC, et al. (2023). Using an artificial intelligence approach to predict the adverse effects and prognosis of tuberculosis. *Diagnostics*. 2023;13(6):1075.
5. Haykin S. *Neural Networks: A Comprehensive Foundation*. 3rd ed., Prentice-Hall, 2007.
6. Orjuela-Canón AD, Jutinico AL, Awad C, Vergara E, Palencia A. Machine learning in the loop for tuberculosis diagnosis support. *Frontiers in Public Health*. 2022;10:876949.
7. Deina C. Aprimorando a tomada de decisão em saúde com aprendizado de máquina em problemas de classificação em dados desbalanceados. 2024.
8. Jannah AW, Al Kindhi B. Optimization of early detection of tuberculosis: Use of multilayer perceptron and extreme learning machine with clinical data. *Jurnal Indonesia Sosial Teknologi*. 2024;5(5).
9. Mohidem NA, Osman M, Muharam FM, Elias SM, Shaharudin R, Hashim Z. Prediction of tuberculosis cases based on sociodemographic and environmental factors in gombak, selangor, malaysia: A comparative assessment of multiple linear regression and artificial neural network models. *The International Journal of Mycobacteriology*. 2021;10(4):442–456.
10. Kanesamoorthy K, Dissanayake MB. Prediction of treatment failure of tuberculosis using support vector machine with genetic algorithm. *The International Journal of Mycobacteriology*. 2021;10(3):279–284.
11. Rodrigues MGA, Sampaio V, Lynn T, Endo PT. A brazilian classified data set for prognosis of tuberculosis, between january 2001 and april 2020.
12. Santos HGD, Nascimento CFD, Izicki R, Duarte YADO, Porto Chiavegatto Filho AD. Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. *Cadernos de Saúde Pública*. 2019;35, e00050818.
13. Talukder MA, Sharmin S, Uddin MA, Islam MM, Aryal S. Mlstd-wsn: machine learning-based intrusion detection using smotetomek in wsns. *International Journal of Information Security*. 2024;23(3):2139–2158.
14. Saouli S, Baarir S, Dutheillet C. Improving sat solver performance through mlp-predicted genetic algorithm parameters. In *International Conference on Integrated Formal Methods*. Springer. 2024: 288–296.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.