

Aprendizagem de máquina na previsão de índices de mortalidade em Unidades de Terapia Intensiva: Uma revisão integrativa da literatura

Machine learning in predicting mortality rates in Intensive Care Units: An integrative review of the literature

Enathanael Ribeiro Soares Mestrando em Ciências da Saúde. Universidade Federal do Cariri (UFCA) – Brasil. <https://orcid.org/0009-0007-3234-8270> enathanael.ribeiro@gmail.com

Maria Angélica Farias Grangeiro Mestranda em Ciências da Saúde. Universidade Federal do Cariri (UFCA) – Brasil. <https://orcid.org/0009-0003-5398-0161> mari_angelica_29@hotmail.com

Ana Heloísa dos Santos Mestranda em Ciências da Saúde. Universidade Federal do Cariri (UFCA) – Brasil. <https://orcid.org/0009-0000-5562-0969> ana.helo0895@gmail.com

Joel Freires de Alencar Arrais Mestrando em Saúde e Sociedade. Universidade do Estado do Rio Grande do Norte (UERN) – Brasil. <https://orcid.org/0000-0002-5127-5309> joel.freires00@gmail.com

Glêbia Alexa Cardoso Doutora em Ciências da Saúde. Universidade Federal da Paraíba (UFPB) – Brasil. <https://orcid.org/0000-0003-4822-1673> glebiacardoso@uern.br

Estelita Lima Cândido Pós-Doutor em Ciências da Saúde. Faculdade de Medicina do ABC (FMABC) – Brasil. <https://orcid.org/0000-0001-9434-2930> estelita.lima@ufca.edu.br

RESUMO

A Unidade de Terapia Intensiva é um ambiente complexo, destinado ao cuidado de pacientes críticos que necessitam de monitoramento constante e intervenções intensivas. A informatização dos prontuários médicos trouxe avanços significativos na gestão de informações, possibilitando a análise de dados em larga escala. Nesse cenário, o aprendizado de máquina (ML) emergiu como uma ferramenta revolucionária para a predição de mortalidade. Realizou-se uma revisão integrativa, abrangendo estudos publicados entre 2020 e 2024 nas bases de dados: Web of Science, LILACS, SciELO e PubMed. Foram analisados 2.627 estudos, com 24 incluídos. A pesquisa focou na aplicação de ML para predição de mortalidade. Ao todo, os estudos usaram dados de 525.081 pacientes provenientes de bases como MIMIC-III, MIMIC-IV e eICU. Os algoritmos mais destacados foram XGBoost, LightGBM, Random Forest e redes neurais (LSTM, GRU), com desempenhos superiores em AUROC, sensibilidade e especificidade. Variáveis-chave incluíram idade, lactato, creatinina, SOFA, SAPS II/III e biomarcadores inflamatórios como IL-6 e LDH. Os modelos de ML demonstraram capacidade de validação de transformar o manejo de pacientes críticos ao proporcionar análises personalizadas e predições mais precisas.

Palavras-chave: Inteligência artificial; aprendizado de máquina supervisionado; Unidade de Terapia Intensiva; hospitalização; previsão.

ABSTRACT

The Intensive Care Unit is a complex environment designed to care for critically ill patients who require constant monitoring and intensive interventions. The computerization of medical records has brought significant advances in information management, enabling large-scale data analysis. In this scenario, machine learning (ML) has emerged as a revolutionary tool for mortality prediction. An integrative review was conducted, covering studies published between 2020 and 2024 in the following databases: Web of Science, LILACS, SciELO, and PubMed. A total of 2,627 studies were analyzed, of which 24 were included. The research focused on the application of ML for mortality prediction. In total, the studies used data from 525,081 patients from databases such as MIMIC-III, MIMIC-IV, and eICU. The most prominent algorithms were XGBoost, LightGBM, Random Forest and neural networks (LSTM, GRU), with superior performance in AUROC, sensitivity and specificity. Key variables included age, lactate, creatinine, SOFA, SAPS II/III and inflammatory biomarkers such as IL-6 and LDH. ML models have demonstrated the validated ability to transform the management of critically ill patients by providing personalized analyses and more accurate predictions.

Keywords: Artificial intelligence; machine learning; Intensive Care Unit; hospitalization; prediction.

Recebido em 19/02/2025. Aprovado em 19/08/2025. Avaliado pelo sistema *double blind peer review*. Publicado conforme normas da ABNT.

<https://doi.org/10.22279/navus.v16.2091>

1 INTRODUÇÃO

A Unidade de Terapia Intensiva (UTI) é um espaço extremamente complexo e dinâmico, destinado ao atendimento de pacientes em estado crítico que requerem vigilância contínua e intervenções intensivas (BACKES; ERDMANN; BÜSCHER, 2015), onde a preocupação com o armazenamento das informações clínicas deve ser redobrada. Nesse sentido, a digitalização dos prontuários médicos representa uma mudança significativa na maneira como as informações dos pacientes e dos serviços são administradas (FARIAS et al., 2011). Essa modernização oferece diversos benefícios que afetam diretamente a eficiência, segurança e qualidade do atendimento em saúde. A informatização dos registros de saúde pode contribuir para a gestão das informações dos pacientes e aprimorar os serviços de saúde, mesmo em nações de baixa e média renda (Numair et al., 2021).

Diante desse contexto, o aprendizado de máquina (ML) surgiu como uma solução inovadora, apresentando técnicas avançadas para a análise de dados e a previsão de resultados significativos, incluindo taxas de mortalidade, readmissões hospitalares e progressão de estados críticos de saúde. Ferramentas como *Gradient Boosting Machines*, redes neurais e algoritmos fundamentados em atenção têm demonstrado um desempenho superior em comparação aos métodos convencionais, como SAPS II e SOFA, permitindo previsões mais acuradas e intervenções fundamentadas em evidências clínicas. (Yu et al., 2024).

Para contextualizar, o aprendizado de máquina, conhecido como machine é um subcampo da ciência da computação que se dedica à análise e desenvolvimento de algoritmos computacionais que aprendem a partir de dados em vez de seguirem instruções pré-definidas (DEO, 2015). A finalidade principal de um modelo de ML é criar um sistema computacional que seja capaz de aprender a partir de um conjunto de dados específico e, ao final desse processo, produzir um modelo para previsões, classificações ou detecções (Jordan; Mitchell, 2015).

Além do impacto técnico, a implementação do aprendizado de máquina nas UTIs tem o potencial de revolucionar o atendimento aos pacientes. Esses modelos conseguem identificar pacientes em alto risco, aprimorar a utilização dos recursos hospitalares e diminuir os custos relacionados ao tratamento intensivo. Além disso, eles tornam a comunicação com os familiares mais eficiente, fornecendo prognósticos mais precisos e embasados em dados concretos. Dessa forma, a tecnologia não só eleva a eficiência, mas também contribui para um atendimento mais humanizado, aperfeiçoando a precisão e a justificativa das decisões médicas (Pessoa et al., 2022).

Apesar dos avanços significativos, o manejo de pacientes em UTIs ainda enfrenta desafios, especialmente na estratificação precoce do risco de mortalidade, limitada por métodos convencionais que por vezes não integram adequadamente dados complexos e variáveis temporais, resultando em previsões imprecisas e atrasos terapêuticos. Este trabalho busca revisar as evidências sobre o uso de aprendizado de máquina na previsão de mortalidade em UTIs, destacando avanços, modelos promissores, variáveis-chave, desafios de implementação e estratégias para integrar essas tecnologias de forma ética e eficaz no contexto clínico.

2 METODOLOGIA

Este estudo é uma revisão integrativa da literatura e obedece às instruções do manual para síntese de evidências do *Joanna Briggs Institute* (JBI) (Aromataris et al., 2024) e do *Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews* (Prisma-ScR) (Tricco et al., 2018), produzida entre os meses de outubro e novembro de 2024. A análise dos estudos incluídos em três bases de dados eletrônicas: *Web of Science*, *Literatura Latino-americana e do Caribe em Ciências da Saúde* (LILACS), *Scientific Electronic Library Online* (SciELO), e *National Library of Medicine* (PubMed).

Em observância ao PICO para pesquisas não-clínicas, objetivou responder a seguinte pergunta norteadora: Quais modelos de aprendizagem de máquina têm sido empregados para prever índices de mortalidade em unidades de terapia intensiva, e como seu desempenho se compara aos métodos tradicionais em termos de acurácia e impacto nas decisões clínicas? Para isso, na formulação da estratégia de busca foram utilizados os seguintes descritores: "Artificial Intelligence", "Machine Learning", "Intensive Care Unit", "Hospitalization", "Prediction" e "Screening, adotando-se como operadores booleanos AND e OR. Todos os descritores estão indexados nas plataformas Descritores em Ciências da Saúde (DeCS) e Medical Subject Heading Terms (MeSH) (Quadro 1).

Quadro 1 - Estratégias de busca

Bases de dados	Operadores booleanos	Estratégias de busca
Web of Science	AND/OR	((("Artificial Intelligence" OR "Machine Learning")) AND ((("Intensive Care Unit" OR "Hospitalization")) AND ((("Prediction" OR " Mass Screening."))
PUBMED	AND/OR	("Artificial Intelligence"OR "Machine Learning") AND ("Intensive Care Units" OR "Hospitalization") AND ("Prediction" OR " Mass Screening.") ("Artificial Intelligence"OR "Machine Learning") AND ("Intensive Care Units" OR "Hospitalization") AND ("Prediction" OR " Mass Screening.") AND ("Mortality")
Scielo	AND/OR	("Inteligência artificial" OR "Aprendizagem de máquina") AND ("Unidade de Terapia Intensiva" OR "Hospitalização") AND ("Predição" OR " Programas de Rastreamento ")
LILACS	AND/OR	("Inteligência Artificial" OR "Aprendizado de Máquina") AND ("Unidade de Terapia Intensiva" OR "Hospitalização") AND ("Predição" OR " Mass Screening ")

Fonte: Dados da Pesquisa, 2025.

Os critérios de inclusão considerados foram os seguintes: a publicação no intervalo de 2020 a 2024; não importando o idioma em que foi publicado nem o local onde o trabalho foi realizado, além de pesquisas baseadas em dados primários. Quanto aos critérios de exclusão: artigos que se mostrassem inconclusivos e/ou estudos repetidos; realizados em modelo animal. Foram excluídos também estudos de revisão e meta análise.

A triagem e classificação dos estudos foi realizado utilizando Rayyan® como ferramenta computacional permitindo a seleção dos trabalhos a partir da análise dos títulos e resumos. Posteriormente realizou-se a leitura na íntegra dos estudos considerando aspectos relevantes, qualidade metodológica e adequação ao foco da pesquisa.

Os estudos selecionados foram distribuídos e organizados na forma de tabela, confeccionada utilizando o Software Microsoft® Word, na versão 2019, com a classificação dos estudos por autor(es), ano de publicação, tamanho da amostra, algoritmo de destaque e principais desfechos.

2 RESULTADOS

A busca nas bases de dados resultou em 2.627 publicações. Após a exclusão de duplicatas (n=514), 2.113 estudos foram considerados para análise inicial. A triagem subsequente, baseada na leitura de títulos e resumos, permitiu uma avaliação preliminar da relevância dos estudos em relação ao objetivo da pesquisa. Foram selecionados 48 artigos para análise do texto na íntegra, dos quais 24 foram excluídos em observância a relevância e acurácia metodológica. Assim, a amostra final foi composta por 24 estudos (Figura 1).

Um total de 525.081 indivíduos participaram das pesquisas (Tabela 1). Todos os estudos versam sobre predição de mortalidade em pacientes críticos internados em Unidades de Terapia Intensiva. Eles exploram a aplicação de algoritmos de aprendizado de máquina para melhorar a precisão preditiva em diferentes condições clínicas.

Os 24 estudos analisados englobaram dados de pacientes oriundos de diferentes bases de dados, como MIMIC-III, MIMIC-IV, eICU, além de bancos de dados regionais. Os indivíduos foram selecionados com base em condições específicas, incluindo insuficiência cardíaca, COVID-19, diabetes tipo 2, sepse, lesão cerebral traumática, câncer colorretal e acidose láctica, com subgrupos representativos destinados à validação interna e externa.

Os algoritmos mais destacados foram XGBoost, LightGBM, Random Forest, *Support Vector Machine* (SVM) e redes neurais como LSTM e GRU. Escalas clínicas frequentemente usadas como variáveis preditoras incluíram SOFA, SAPS II/III, APACHE II e escores como Charlson e Elixhauser. Algumas variáveis foram usadas de forma relevante nos estudos, como idade, lactato, ureia, troponina, plaquetas, pressão arterial, glicose e biomarcadores inflamatórios como IL-6, LDH e CRP.

Tabela 1 - Caracterização dos 24 estudos analisados

(continua)

Autor Ano	n	Base de Dados Médica ou Local do estudo	Desenho do Estudo	Algoritmos de destaque
AHMAD et al. 2021	4.000	MIMIC-III	Estudo retrospectivo com aplicação de técnicas de aprendizado de máquina.	Support Vector Machine
ALDRAIMLI et al. 2022	48.546	MIMIC-III	Estudo comparativo de classificação binária.	XGBoost
BARCHITTA et al. 2021	3.782	Rede SPIN-UTI na Itália	Estudo multicêntrico retrospectivo.	SVM
CHEN et al. 2022	18.094	MIMIC-III	Estudo retrospectivo com aprendizado profundo.	ATCN
CHEN et al. 2023	2.803	MIMIC-IV e eICU	Estudo retrospectivo com validação externa.	XGBoost
CHIU et al. 2022	6.699	MIMIC-III	Estudo retrospectivo com aplicação de ML.	Stacking Ensemble Model
DANAY et al. 2024	2.331	eICU	Coorte retrospectivo.	XGBoost
GARCÍA-GALLO et al. 2020	5.650	MIMIC-III	Coorte retrospectiva.	SGB
GEORGE et al. 2021	4.334	MIMIC-III	Coorte retrospectiva.	DNN
GUO et al. 2023	969	eICU, MIMIC-IV e Hospital de Wuhan.	Estudo retrospectivo multicêntrico.	Ensemble Model (Random Forest, LightGBM e XGBoost)
HU et al. 2024	14.222	MIMIC-III	Estudo retrospectivo com validação cruzada.	Bagging e AdaBoost
KANG et al. 2021	219.306	MIMIC-III e eICU	Desenvolvimento e validação de modelo.	DNN com LSTM
KIM; KWON, BAEK 2021	16.940	Hospitais na Coreia do Sul	Estudo retrospectivo comparativo.	XGBoost
NISTAL-NUÑO et al. 2020	9.893	MIMIC-III	Comparação de modelos.	XGBoost
PANG et al. 2022	67.748	MIMIC-IV	Estudo comparativo.	XGBoost
NA PATTALUNG; INGVIIYA; CHAICHULLE 2021	75.685	MIMIC-III, MIMIC-IV e eICU	Estudo comparativo.	GRU
PATTHARANITIMA et al. 2021	1.919	MIMIC-III	Estudo retrospectivo.	Random Forest
PEZOULAS et al. 2022	214	Hospital "Sotiria" para doenças do tórax, Atenas	Modelagem explicável de IA.	GBDT
SAKAGIANNI et al. 2023	373	Hospital Terciário na Grécia	Estudo observacional retrospectivo.	Random Forest

TU et al. 2023	2.260	Hospital Chi Mei Medical Center, Taiwan	Estudo retrospectivo comparando modelos de ML.	Random Forest e LightGBM (conclusão)
Autor Ano	n	Base de Dados Médica ou Local do estudo	Desenho do Estudo	Algoritmos de destaque
YE et al. 2023	5.247	MIMIC-IV eICU	Estudo retrospectivo com aplicação de técnicas de aprendizado de máquina.	GBDT
YU et al. 2024	9.118	MIMIC-III	Estudo retrospectivo com aplicação de técnicas de aprendizado de máquina.	LightGBM
ZHAI et al. 2024	1.668	ICU Open Source (China)	Estudo retrospectivo.	GBDT
ZHOU et al. 2024	3.280	MIMIC-III, MIMIC-IV e eICU	Coorte retrospectiva com validação cruzada.	XGBoost

LSTM = Long Short-Term Memory; SHAP = SHapley Additive exPlanations; ML = Machine Learning; LDA = Linear Discriminant Analysis; LightGBM = Light Gradient Boosting Machine; GBDT = Gradient Boosting Decision Tree; DNN = Deep Neural Network; XGBoost = Extreme Gradient Boosting; SVM = Support Vector Machines; SGB = Stochastic Gradient Boosting; GRU = Gated Recurrent Unit; ATCN = Attention-based Temporal Convolution Network.
Fonte: Dados da pesquisa, 2025.

4 DISCUSSÃO

A aplicação de modelos de aprendizado de máquina na predição de mortalidade, especialmente na UTI, parece refletir a crescente busca por maior personalização e cuidado do paciente criticamente enfermo. Inúmeros algoritmos foram analisados para variadas populações e cenários clínicos, com ênfase nos modelos fundamentados em ML, como *Gradient Boosting Machines*, Random Forest e suas adaptações, além de métodos mais sofisticados, como Redes Neurais Recorrentes (RNN) e modelos explicáveis.

A acurácia dos modelos preditivos adotados nos estudos foi avaliada e comparada com métodos tradicionais através da AUROC. A curva ROC (*Receiver Operating Characteristic*) compara o desempenho de modelos classificatórios representando a ligação entre as taxas de verdadeiros positivo (sensibilidade) e falsos positivos em diferentes situações preditivas (Müller; Guido, 2016). A AUC (*Área Under the Curve*) quantifica essa curva, fornecendo um valor numérico que expressa a capacidade do modelo de distinguir entre as classes (Hosmer; Lemeshow; Sturdivant, 2013), tornando-a especialmente útil para cenários com classes desbalanceadas, como na predição de mortalidade hospitalar. A interpretação dessa métrica pode ser vista na Tabela 2.

Tabela 2 - Interpretação/Classificação dos valores de AUC

AUROC	Interpretação	Desempenho do Modelo
0,5-0,6	Sem poder de discriminação	Péssimo
0,6-0,7	Baixa capacidade preditiva	Ruim
0,7-0,8	Moderada capacidade preditiva	Pobre
0,8-0,9	Boa capacidade preditiva	Bom
>0,9	Excelente capacidade de discriminação entre as classes	Excelente

Fonte: Adaptado de Polo e Miot (2020).

A adesão de técnicas de explicabilidade como SHAP (*Shapley Additive exPlanations*) e LIME (*Local Interpretable Model-Agnostic Explanations*) é um elemento comum adotado pelos estudos, facilitando a identificação de variáveis fundamentais e tornando as predições mais compreensíveis para a aplicação na clínica, como em Na Pattalung et al. (2021), Zhai et al. (2024), Pang et al. (2022) e Nistal-Nuño (2022), destaque para Pezoulas et al. (2022) que combinaram redes bayesianas dinâmicas com SHAP para prever mortalidade em pacientes com COVID-19, destacando citocinas inflamatórias (IL-6, IL-8 e TNF) como preditores críticos. Essa abordagem multimodal ofereceu percepções valiosas sobre a dinâmica biológica e clínica.

García-Gallo et al. (2020) e Yu et al. (2024) utilizaram o SGB e o LightGBM, respectivamente, para pacientes com sepse. Ambos os estudos reportaram alta acurácia (AUC acima de 0,80), com variáveis como idade, níveis de bilirrubina e lactato desempenhando papéis centrais. Essa abordagem superou métodos tradicionais como SAPS II e SOFA, evidenciando a superioridade dos modelos de ML na análise preditiva, sendo no estudo de García-Gallo et al. (2020) capaz de prever a mortalidade em 1 ano, usando dados das primeiras 24 horas após a admissão na UTI. Nessa linha de raciocínio, Zhou et al. (2024) empregaram XGBoost para mortalidade associada a coagulopatia por sepse, destacando explicações interpretáveis com SHAP, o que fortalece a integração clínica.

Ademais, Pang et al. (2022) e Nistal-Nuño (2022) também utilizaram o XGBoost, que obteve um AUROC superior a 0,91, ultrapassando algoritmos como SVM e Árvore de Decisão, além de SOFA e SAPS. Em ambos os estudos, a aplicação

do SHAP para interpretar as previsões elevou a confiança nos resultados, destacando variáveis importantes, como a Escala de Coma de Glasgow (GCS) e a administração de vasopressores.

Essa abordagem prática também foi reforçada por Pattharanitima et al. (2021), cujo modelo de Árvore de decisão, com AUROC de 0,83, foi implementado para prever mortalidade em pacientes com acidose láctica, superando o SAPS II.

Kim et al. (2021) mostraram que o XGBoost é eficaz na previsão de mortalidade em 30 dias para pacientes sob ventilação mecânica, com AUC de 0,79, destacando variáveis como o escore APACHE II e o Índice de Comorbidade de Charlson.

Os estudos de Chiu et al. (2022) e Kang et al. (2022) exploraram populações mais amplas, como pacientes com insuficiência cardíaca e UTIs multicêntricas. O Stacking Ensemble Model obteve AUC de 0,82 para previsões de curto prazo, tendo precisão geral de 95,25% na previsão de mortalidade em 3 dias, enquanto o LSTM com dois níveis de atenção AUC de 0.855 na validação externa, proporcionando alta precisão e interpretabilidade, indicando que o uso de dados longitudinais pode melhorar a precisão.

Pacientes com condições específicas, como câncer colorretal (GUO et al., 2023) e lesão cerebral traumática (TU et al., 2023), beneficiaram-se de abordagens personalizadas, como LightGBM e aprendizado por agrupamento, que integraram dados clínicos e marcadores específicos. O conjunto combinando Random Forest, LightGBM e XGBoost, superou modelos individuais, com AUC de 0.86 no treinamento e 0.73 na validação externa, fornecendo suporte a decisões clínicas críticas, com maior confiabilidade. O LightGBM teve desempenho significativamente superior ao APACHE II (AUC = 0,872) e SOFA (AUC = 0,853).

Aldraimli et al. (2022), sob este ponto de vista, utilizaram a árvore de decisão em conjunto com esta técnica para tratar o desequilíbrio de dados, obtendo uma AUROC de 0,94. Este estudo demonstrou a importância da preparação adequada dos dados para maximizar o desempenho do modelo, particularmente em populações onde as classes são altamente desproporcionais.

Modelos baseados em redes neurais também ganharam destaque. George et al. (2021) utilizaram *Deep Neural Networks* (DNN) para prever mortalidade em paciente com uso de ventilação mecânica prolongada, superando significativamente o ProVent Score (AUC 0,76 X AUC 0,63, respectivamente), principalmente para mortalidade em 12 meses. Na Pattalung et al. (2021) demonstraram o desempenho robusto de GRU (AUC > 0,80) ao analisar séries temporais, reforçando o papel de redes otimizadas em cenários críticos.

A questão do desbalanceamento de dados, frequentemente observada em populações de UTI (onde a quantidade de sobreviventes é superior à de óbitos), foi tratada de maneiras variadas. Ahmad et al. (2021) e Aldraimli et al. (2022) utilizaram o SMOTE como estratégia para aumentar a representação da classe menos favorecida (mortalidade), resultando em uma melhoria significativa na sensibilidade dos modelos. Em contrapartida, estudos como o de García-Gallo et al. (2020) não incluíram técnicas de balanceamento, indicando que a falta desse tipo de pré-processamento pode restringir a habilidade de identificar com precisão pacientes em alto risco. Essa divergência ressalta a importância de métodos de balanceamento para otimizar a eficácia das previsões.

No que diz respeito à generalização dos modelos, Chen et al. (2022) enfatizam a relevância da validação externa em bancos de dados como MIMIC-IV e eICU-CRD, reforçando as conclusões de Guo et al. (2023) e Pezoulas et al. (2022), que asseguram a robustez e a aplicabilidade clínica desses modelos.

Pesquisas que se restringem a uma única base de dados podem apresentar limitações em termos de aplicabilidade prática, especialmente em populações variadas ou em contextos regionais diferentes.

Ye et al. (2023) utilizaram o eICU-CRD na validação de seus modelos. Essa abordagem promoveu uma maior generalização dos resultados, ampliando a relevância clínica. A variedade de populações e condições clínicas analisadas, incluindo ventilação mecânica prolongada, acidose láctica e insuficiência cardíaca, evidenciou a adaptabilidade dos algoritmos em distintos cenários de UTI.

Os principais resultados obtidos pelos modelos indicam um impacto considerável na prática clínica. A habilidade de prever a mortalidade em curto e longo prazo (Barchitta et al., 2021; Hu et al., 2024) pode auxiliar na tomada de decisões sobre a alocação de recursos, priorização de cuidados e na realização de intervenções precoces. Ferramentas online, como a desenvolvida por Chen et al. (2023), mostram de que maneira o aprendizado de máquina pode ser incorporado aos processos hospitalares, no caso para pacientes diabéticos.

Essa integração prática contrasta com estudos puramente exploratórios, como o de Sakagianni et al. (2023), sugerindo que a translação do modelo para a prática clínica deve ser um objetivo prioritário nos próximos anos.

Apesar dos avanços, persistem desafios como a necessidade de maior padronização entre os estudos e maior adoção de validações multicêntricas. Além disso, a inclusão de variáveis dinâmicas e temporais, como demonstrado por Danay et al. (2024), deve ser priorizada para capturar a evolução clínica dos pacientes.

5 CONSIDERAÇÕES FINAIS

Os modelos de ML demonstram potencial revolucionário na previsão de mortalidade, incluindo os cenários da UTI, fornecendo precisão superior e insights interpretáveis. A junção de biomarcadores e dados clínicos relevantes possibilitam a personalização do cuidado, eficácia e decisões baseadas em evidências palpáveis. Esses modelos apresentaram melhores resultados superando sistemas tradicionais como SAPS II e SOFA.

Destaque para XGBoost, LightGBM e RNN, sendo excelentes na integração de variáveis clínicas e temporais, permitindo análises mais precisas. SHAP e LIME foram usadas com maestria como ferramenta explicativa, facilitando a aceitação clínica. Outro sim, algoritmos dinâmicos como ATCN, rede neural baseado na atenção particularmente eficaz em séries históricas identificando mudanças fisiológicas importantes.

Apesar do progresso, desafios como a necessidade de padronização metodológica, processo de validações externas multicêntricas e o desequilíbrio de dados ainda limitam a aplicação universal dos modelos. A inclusão de variáveis críticas, como idade, escore SOFA, lactato e creatinina, reforça a importância dessas abordagens, enquanto modelos eficientes, como o LightGBM, garantem aplicabilidade prática. No geral, a integração do ML nos sistemas hospitalares tem o potencial de transformar os cuidados críticos, fornecendo previsões rápidas, precisas e personalizadas, promovendo intervenções mais precoces e melhores resultados para pacientes gravemente enfermos.

REFERÊNCIAS

- AHMAD, Fawad; AYUB, Huma; LIAQAT, Rehan; KHAN, Akhyar Ali; NAWAZ, Ali; YOUNIS, Babar. **Mortality Prediction in ICU Patients Using Machine Learning Models**. In: INTERNATIONAL BHURBAN CONFERENCE ON APPLIED SCIENCES AND TECHNOLOGIES (IBCAST), 2021, Islamabad, Pakistan. Anais... Islamabad: IEEE, 2021. p. 372-376. DOI: 10.1109/IBCAST51254.2021.9393012.
- ALDRAIMLI, Mahmoud; NAZYROVA, Nodira; DJUMANOV, Abdumalik; SOBIROV, Ikboljon; CHAUSSALET, Thierry. A Comparative Machine Learning Modelling Approach for Patients' Mortality Prediction in Hospital Intensive Care Unit. **BioInfoMed 2020**, v.374, n.1, p.15-31, 2022.
- AROMATARIS, Edoardo; LOCKWOOD, Craig; PORRITT, Kylie; PILLA, Bianca; JORDAN, Zoe (eds.). **JB1 Manual for Evidence Synthesis**. JBI, 2024. Disponível em: <https://synthesismanual.jbi.global>. DOI: 10.46658/JBIMES-24-01.
- BACKES, Marli Terezinha Stein; ERDMANN, Alacoque Lorenzini; BÜSCHER, Andreas. O ambiente vivo, dinâmico e complexo de cuidados em Unidade de Terapia Intensiva. **Revista Latino-Americana de Enfermagem**, v. 23, n. 3, p. 411-418, 2015.
- BARCHITTA, Martina; MAUGERI, Andrea; FAVARA, Giuliana; RIELA, Paolo Marco; GALLO, Giovanni; MURA, Ida; AGODI, Antonella. Early Prediction of Seven-Day Mortality in Intensive Care Unit Using a Machine Learning Model: Results from the SPIN-UTI Project. **Journal of Clinical Medicine**, v. 10, n. 5, p. 992, 2021.
- CHEN, Yu-wen; LI, Yu-jie; DENG, Peng; YANG, Zhi-yong; ZHONG, Kun-hua; ZHANG, Li-ge; CHEN, Yang; ZHI, Hong-yu; HU, Xiao-yan; GU, Jian-teng; NING, Jiao-lin; LU, Kai-zhi; ZHANG, Ju; XIA, Zheng-yuan; QIN, Xiao-lin; YI, Bin. Learning to predict in-hospital mortality risk in the intensive care unit with attention-based temporal convolution network. **BMC Anesthesiology**, v. 22, n. 1, p. 119, 2022.
- CHEN, Zijun; LI, Tingming; GUO, Sheng; ZENG, Deli; WANG, Kai. Machine learning-based in-hospital mortality risk prediction tool for intensive care unit patients with heart failure. **Frontiers in Cardiovascular Medicine**, v. 10, p. 1119699, 2023.
- CHIU, Chih-Chou; WU, Chung-Min; CHIEN, Te-Nien; KAO, Ling-Jing; LI, Chengcheng; JIANG, Han-Ling. Applying an Improved Stacking Ensemble Model to Predict the Mortality of ICU Patients with Heart Failure. **Journal of Clinical Medicine**, v. 11, n. 21, p. 6460, 2022.
- DANAY, Lihi; RAMON-GONEN, Roni; GORODETSKI, Maria; SCHWARTZ, David G. Evaluating the effectiveness of a sliding window technique in machine learning models for mortality prediction in ICU cardiac arrest patients. **International Journal of Medical Informatics**, v. 191, p. 105565, 2024.
- DEO, Rahul C. Machine Learning in Medicine. **Circulation**. v.132, n.20, p.192030, 2015.

FARIAS, Josivania Silva; GUIMARÃES, Tomás de Aquino; VARGAS, Eduardo Raupp de; ALBUQUERQUE, Pedro Henrique Melo. Adoção de prontuário eletrônico do paciente em hospitais universitários de Brasil e Espanha: a percepção de profissionais de saúde. **Revista de Administração Pública**, v. 45, n. 5, p. 1303-1326, 2011.

GARCÍA-GALLO, Juan E.; FONSECA-RUIZ, Nelson J.; CELI, Leo A.; DUITAMA-MUÑOZ, Juan F. A machine learning-based model for 1-year mortality prediction in patients admitted to an Intensive Care Unit with a diagnosis of sepsis. **Medicina Intensiva**, v. 44, n. 3, p. 160-170, 2020.

GEORGE, Naomi; MOSELEY, Edward; EBER, Rene; SIU, Jennifer; SAMUEL, Mathew; YAM, Jonathan; HUANG, Kexin; CELI, Leo Anthony; LINDVALL, Charlotta. Deep learning to predict long-term mortality in patients requiring 7 days of mechanical ventilation. **PLOS ONE**, v. 16, n. 6, p. e0253443, 2021.

GUO, Chunxia; PAN, Jun; TIAN, Shan; GAO, Yuanjun. Using machine learning algorithms to predict 28-day mortality in critically ill elderly patients with colorectal cancer. **Journal of International Medical Research**, v. 51, n. 11, p. 03000605231198725, 2023.

HOSMER, David W. Jr.; LEMESHOW, Stanley; STURDIVANT, Rodney X. **Applied Logistic Regression**. 3. ed. Hoboken: John Wiley & Sons, Inc., 2013.

HU, Tung-Lai; CHAO, Chuang-Min; WU, Chien-Chih; CHIEN, Te-Nien; LI, Chengcheng. Machine Learning-Based Predictions of Mortality and Readmission in Type 2 Diabetes Patients in the ICU. **Applied Sciences**, v. 14, n. 18, p. 8443, 2024.

JORDAN, Michael I.; MITCHELL, Tom M. Machine Learning: Trends, Perspectives, and Prospects. **Science**, v. 349, n. 6245, p. 255-260, 2015.

KANG, Yanni; JIA, Xiaoyu; WANG, Kaifei; HU, Yiyi; GUO, Jianying; CONG, Lin; LI, Xiang; XIE, Guotong. A Clinically Practical and Interpretable Deep Model for ICU Mortality Prediction with External Validation. **AMIA Annual Symposium Proceedings**, v. 2020, p. 629-637, 2021.

KIM, Jong Ho; KWON, Young Suk; BAEK, Moon Seong. Machine Learning Models to Predict 30-Day Mortality in Mechanically Ventilated Patients. **Journal of Clinical Medicine**, v. 10, n. 10, p. 2172, 2021.

MUELLER, C.; GUIDO, S. Introduction to Machine Learning with Python: A guide for data scientist. [S.l.]: O'Reilly Media, 2016.

MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. [S.l.]: O'Reilly Media, 2016.

NA PATTALUNG, Thanakron; INGVIYA, Thammasin; CHAICHULEE, Sitthichok. Feature Explanations in Recurrent Neural Networks for Predicting Risk of Mortality in Intensive Care Patients. **Journal of Personalized Medicine**, v. 11, n. 9, p. 934, 2021.

NISTAL-NUÑO, Beatriz. Developing machine learning models for prediction of mortality in the medical intensive care unit. **Computer Methods and Programs in Biomedicine**, v. 216, p. 106663, 2022.

NUMAIR, Tarek; HARRELL, Daniel Toshio; HUY, Nguyen Tien; NISHIMOTO, Futoshi; MUTHIANI, Yvonne; NZOU, Samson Muuo; LASAPHONH, Angkhana; PALAMA, Khomsonerasinh; PONGVONGSA, Tiengkham; MOJI, Kazuhiko; HIRAYAMA, Kenji; KANEKO, Satoshi. Barriers to the Digitization of Health Information: A Qualitative and Quantitative Study in Kenya and Lao PDR Using a Cloud-Based Maternal and Child Registration System. **International Journal of Environmental Research and Public Health**, v. 18, n. 12, p. 6196, 2021.

PANG, Ke; LI, Liang; OUYANG, Wen; LIU, Xing; TANG, Yongzhong. Establishment of ICU Mortality Risk Prediction Models with Machine Learning Algorithm Using MIMIC-IV Database. **Diagnostics**, v. 12, n. 5, p. 1068, 24 abr. 2022.

PATTHARANITIMA, Pattharawin; THONGPRAYOON, Charat; KAWEPUT, Wisit; QURESHI, Fawad; QURESHI, Fahad; PETNAK, Tananchai; SRIVALI, Narat; GEMBILLO, Guido; O'CORRAGAIN, Oisín A.; CHESDACHAI, Supavit; VALLABHAJOSYULA, Saraschandra; GURU, Pramod K.; MAO, Michael A.; GAROVIC, Vesna D.; DILLON, John J.; CHEUNGPASITPORN, Wisit. Machine Learning Prediction Models for Mortality in Intensive Care Unit Patients with Lactic Acidosis. **Journal of Clinical Medicine**, v. 10, n. 21, p. 5021, 2021.

PESSOA, Stela Mares Brasileiro; OLIVEIRA, Bianca Silva de Sousa; SANTOS, Wendy Gomes dos; OLIVEIRA, Augusto Novais Macedo; CAMARGO, Marianne Silveira; MATOS, Douglas Leandro Aparecido Barbosa de; SILVA, Miquéias Martins Lima; MEDEIROS, Carolina Cintra de Queiroz; COELHO, Cláudia Soares de Sousa; ANDRADE NETO, José de Souza; MISTRO, Sóstenes. Predição de choque séptico e hipovolêmico em pacientes de unidade de terapia intensiva com o uso de machine learning. **Revista Brasileira de Terapia Intensiva**, v. 34, n. 4, p. 477-483, 2022.

PEZOULAS, Vasileios C.; TSIORIS, Konstantinos; ZERVAKIS, Michalis; KONITSIOTIS, Spyridon; FOTIADIS, Dimitrios I. A Multimodal Approach for the Risk Prediction of Intensive Care and Mortality in Patients with COVID-19. **Diagnostics (Basel)**, v. 12, n. 1, p. 1-28, 2021.

POLO, Tatiana Cristina Figueira; MIOT, Hélio Amante. Aplicações da curva ROC em estudos clínicos e experimentais. **Jornal Vascular Brasileiro**, v. 19, p. e20200186, 2020.

SAKAGIANNI, Aikaterini; KOUFOPOULOU, Christina; VERYKIOS, Vassilios; LOUPELIS, Evangelos; KALLES, Dimitrios; FERETZAKIS, Georgios. **Prediction of COVID-19 Mortality in the Intensive Care Unit Using Machine Learning**. In: Studies in Health Technology and Informatics. [S.l.]: IOS Press, 2023. p. 536-540.

TRICCO, Andrea C.; LILLIE, Erin; ZARIN, Wasifa; O'BRIEN, Kelly K.; COLQUHOUN, Heather; LEVAC, Danielle; MOHER, David; PETERS, Micah D. J.; HORSLEY, Tanya; WEEKS, Laura; HEMPEL, Susanne; ATKINS, Dianne; AMENTA, William; MUNIYANDI, Muthu; NGUYEN, Phuong; SIVARAMAN, Meenakshi; STRAUS,

Sharon E. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. **Annals of Internal Medicine**, v. 169, n. 7, p. 467-473, 2018.

TU, Kuan-Chi; TAU, Eric Nyam Tee; CHEN, Nai-Ching; CHANG, Ming-Chuan; YU, Tzu-Chieh; WANG, Che-Chuan; LIU, Chung-Feng; KUO, Ching-Lung. Machine Learning Algorithm Predicts Mortality Risk in Intensive Care Unit for Patients with Traumatic Brain Injury. **Diagnostics**, v. 13, n. 18, p. 3016, 2023.

YE, Zixiang; AN, Shuoyan; GAO, Yanxiang; XIE, Enmin; ZHAO, Xuecheng; GUO, Ziyu; LI, Yike; SHEN, Nan; REN, Jingyi; ZHENG, Jingang. The prediction of in-hospital mortality in chronic kidney disease patients with coronary artery disease using machine learning models. **European Journal of Medical Research**, v. 28, n. 1, p. 33, 2023.

YU, Zhijiang; LI, Hexin; PISHGAR, Maryam; ASHRAFI, Negin; ALAEI, Kiarash. Prediction of 30-day mortality for ICU patients with Sepsis-3. **BMC Medical Informatics and Decision Making**, v. 24, n. 1, p. 223, 2024.

ZHAI, Yihai; LAN, Danxiu; LV, Siying; MO, Liqin. Interpretability-based machine learning for predicting the risk of death from pulmonary inflammation in Chinese intensive care unit patients. **Frontiers in Medicine**, v. 11, p. 1399527, 2024.

ZHOU, Shu; LU, Zongqing; LIU, Yu; WANG, Minjie; ZHOU, Wuming; CUI, Xuanxuan; ZHANG, Jin; XIAO, Wenyan; HUA, Tianfeng; ZHU, Huaqing; YANG, Min. Interpretable machine learning model for early prediction of 28-day mortality in ICU patients with sepsis-induced coagulopathy: development and validation. **European Journal of Medical Research**, v. 29, n. 1, p. 14, 2024.