

16º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2025

DIAGNÓSTICO PRECOCE DE SEPSE POR MEIO DE REDES NEURAIS ARTIFICIAIS

DIOGO L. TOPORCOV¹, ANA PAULA A. C. E SHIGUEMORI², ELCIO H. SHIGUEMORI³

¹ Graduando em Tecnologia de Análise e Desenvolvimento de Sistemas, Bolsista PIBIFSP, IFSP, Campus Jacareí, diogo.toporcov@aluno.ifsp.edu.br.

² Professor do IFSP, Campus Jacareí, anapaula.acs@ifsp.edu.br.

³ Dr. Pesquisador do DCTA, elcioehs@fab.mil.br

Área de conhecimento (Tabela CNPq): 1.03.01.03-8 Análise de Algoritmos e Complexidade de Computação

RESUMO: A sepsé é uma condição médica grave, caracterizada por uma resposta inflamatória desregulada a uma infecção, que pode evoluir rapidamente para falência múltipla de órgãos e óbito. Os sintomas iniciais, como febre e alterações na frequência cardíaca, apresentam caráter inespecífico e podem se confundir com outras condições clínicas, dificultando o diagnóstico. A inteligência artificial (IA), especialmente com Redes Neurais Artificiais (RNAs), têm avançado na medicina, aplicando-se a diagnósticos de tumores, análise de imagens e previsão de complicações clínicas. Este projeto propõe o desenvolvimento de um modelo preditivo baseado em RNAs para identificar sinais precoces de sepsé, possibilitando intervenções clínicas eficazes. Utilizando dados do *Kaggle*, MIMIC-III e do Hospital das Clínicas da USP, o modelo investigará variáveis clínicas críticas associadas à progressão da sepsé. Os resultados indicaram que o modelo obteve boa performance, com alta acurácia e capacidade de discriminação entre classes. No entanto, desafios como o desequilíbrio entre as classes de pacientes e a distorção causada pelos dados ausentes e preenchidos impactaram a precisão em simulações em tempo real. Esses resultados demonstram o potencial de IA no diagnóstico precoce de sepsé, mas destacam necessidades de melhorias no tratamento dos dados faltantes e na adaptação do modelo para cenários clínicos reais.

PALAVRAS-CHAVE: Sepsé; Diagnóstico precoce; Redes Neurais Artificiais (RNA); Inteligência Artificial; Dados clínicos; Modelos preditivos

EARLY DIAGNOSIS OF SEPSIS WITH ARTIFICIAL NEURAL NETWORKS

ABSTRACT: Sepsis is a severe medical condition characterized by a dysregulated inflammatory response to infection, which can rapidly progress to multiple organ failure and death. Initial symptoms, such as fever and changes in heart rate, are nonspecific and may overlap with other clinical conditions, making early diagnosis challenging. Artificial Intelligence (AI), particularly through Artificial Neural Networks (ANNs), has made significant strides in medicine, being applied to tumor diagnosis, image analysis, and prediction of clinical complications. This project proposes the development of a predictive model based on ANNs to identify early signs of sepsis, enabling effective clinical interventions. Using data from Kaggle, MIMIC-III, and the Hospital das Clínicas of USP, the model will investigate critical clinical variables associated with the progression of sepsis. The results indicated that the model performed well, with high accuracy and good discriminatory power between classes. However, challenges such as class imbalance and the distortion caused by missing and imputed data impacted the accuracy in real-time simulations. These results demonstrate the potential of AI in the early diagnosis of sepsis, but also highlight the need for improvements in handling missing data and adapting the model to real clinical scenarios.

KEYWORDS: Sepsis; Early diagnosis; Artificial Neural Networks (ANN); Artificial Intelligence; Clinical data; Predictive models

INTRODUÇÃO

A sepse é uma condição médica grave, caracterizada por uma resposta inflamatória descontrolada a uma infecção, com potencial para rápida evolução fatal. Considerada um desafio mundial na área da saúde, apresenta alta taxa de morbidade e mortalidade. A sepse é definida como “resposta desregulada do organismo a uma infecção, resultando em disfunção orgânica e risco de vida” (Gojak et al., 2022; Manual MSD, 2024). Segundo a Organização Mundial da Saúde (OMS), cerca de 11 milhões de mortes anuais resultam de casos de sepse, agravados pela progressão rápida da doença e pela dificuldade em revertê-la quando o tratamento é iniciado tardiamente (Moura et al., 2017). Por isso, é essencial a identificação precoce e a ação imediata de profissionais da saúde para reduzir a mortalidade (Silva, 2022; Pereira Junior et al., 1998).

A inteligência artificial (IA), especialmente por meio de redes neurais artificiais (RNAs), vem avançando na medicina, sendo aplicada em diagnósticos, análise de imagens e previsão de complicações (Gojak et al., 2022; Pais et al., 2023). Este projeto propõe desenvolver um modelo de IA baseado em RNAs para identificar sinais precoces de sepse, utilizando dados do *Kaggle*, uma plataforma *online* que hospeda competições de ciência de dados e aprendizado de máquina, oferecendo conjuntos de dados públicos e notebooks para codificação colaborativa, promovendo a interação entre cientistas de dados e pesquisadores (Bojer et al., 2021), MIMIC-III (Johnson et al., 2016) e do Hospital das Clínicas da USP, em parceria com o ITA e o IFSP. O modelo investigará variáveis clínicas críticas associadas à progressão da doença. As redes neurais, com sua capacidade de analisar dados sequenciais como o histórico médico de pacientes internados, permitem identificar a evolução da sepse e a deterioração do quadro clínico, favorecendo decisões mais informadas e uma resposta mais eficaz em situações críticas (Sherstinsky, 2020; GARG; Bansal, 2023). Espera-se que este sistema funcione como uma ferramenta de apoio à decisão clínica, contribuindo para a redução da mortalidade e maior precisão no diagnóstico (Sharan et al., 2023).

MATERIAL E MÉTODOS

Esta seção apresenta a estrutura metodológica, os recursos utilizados e as técnicas aplicadas para o desenvolvimento do modelo. O foco está na organização dos dados, nas estratégias de pré-processamento, na implementação das arquiteturas de redes neurais artificiais e na avaliação do desempenho obtido com os experimentos.

1. Conjuntos de Dados

O estudo utilizou inicialmente o conjunto público disponível no *Kaggle*, composto por registros clínicos de pacientes monitorados hora a hora. Considerou-se também o uso do conjunto de dados clínicos MIMIC-III (Johnson et al., 2016) como fonte complementar, no entanto, esse conjunto não possuía rótulos pré-definidos indicando casos positivos ou negativos de sepse, o que exigiria um estudo adicional voltado à aplicação de técnicas de classificação. Diante dessa limitação, optou-se temporariamente pelo conjunto de dados do Kaggle, que, embora menor, já disponibiliza os rótulos prontos.

2. Pré-processamento e Preparação dos Dados

Os registros clínicos analisados incluíam uma série de variáveis, como sinais vitais (frequência cardíaca, saturação de oxigênio, temperatura, pressão arterial sistólica, média e diastólica, frequência respiratória), dados laboratoriais (como gases sanguíneos, bilirrubinas, glicose, lactato, creatinina, troponina, entre outros), informações sobre a composição sanguínea (hematócrito, hemoglobina, plaquetas, leucócitos), parâmetros de ventilação (pH, PaCO₂, EtCO₂), e dados demográficos (idade, gênero), além de informações relacionadas à internação (tempo de UTI, tempo de admissão no hospital, rótulos se o paciente apresentava sepse, etc.). Entretanto, esses registros apresentavam muitos valores ausentes, especialmente nas variáveis laboratoriais. Isso ocorre porque esses exames não são

realizados continuamente, mas apenas quando solicitados pelo médico, conforme a necessidade clínica. Em contraste, sinais vitais como frequência cardíaca, temperatura e pressão arterial são coletados automaticamente por equipamentos de monitoramento, resultando em medições mais frequentes e regulares (Queiroz, 2019). Já os exames laboratoriais aparecem apenas nos horários específicos em que foram coletados, o que explica os longos períodos sem dados. Inicialmente, foram removidas as variáveis com ausência excessiva de dados ou de baixa relevância clínica, a fim de assegurar a integridade e a pertinência das informações utilizadas, seguindo com a interpolação temporal para preencher lacunas, preservando a sequência cronológica dos eventos clínicos de cada paciente. Por fim, houve uma padronização temporal por meio da aplicação de *padding*, uma técnica que consiste em adicionar valores nulos nas sequências mais curtas para uniformizar o comprimento das sequências para um intervalo fixo. Por fim, os dados foram organizados em tensores tridimensionais, possibilitando a captura e a análise de padrões sequenciais na evolução clínica dos pacientes (Lipton et al., 2016).

3. Divisão dos Dados

Para a divisão dos dados, foi realizada uma seleção inicial de 5 pacientes positivos (que apresentaram sepse em algum ponto da sequência) e 5 pacientes negativos (que nunca apresentaram sepse ao longo de toda a sequência), de forma que esses pacientes não fossem incluídos nos conjuntos de treinamento, validação ou teste. Após essa seleção, os dados foram então divididos nos conjuntos de treinamento, validação e teste, sendo que 70% dos dados foram utilizados para treinamento, 15% para validação e 15% para teste.

4. Arquitetura do Modelo

A arquitetura inicial adotada foi uma Rede Neural Convolutacional Temporal (TCN), composta por blocos convolucionais dilatados e conexões residuais (Bai et al., 2018). O modelo foi desenvolvido para capturar dependências temporais de longo alcance, as quais são cruciais para a modelagem da progressão da sepse. A escolha da TCN se baseia em sua capacidade comprovada de lidar com sequências temporais longas, ao permitir a expansão do campo receptivo de maneira eficiente (Bai et al., 2018), além de possibilitar a modelagem de dados temporais com alta acurácia em problemas de previsão, como é o caso da detecção de sepse (Lea et al., 2017). Essa estrutura foi implementada em *Python*, utilizando o *framework PyTorch* (Paszke et al., 2019), com otimização para aceleração em GPU a fim de reduzir o tempo de treinamento.

5. Treinamento e Ajuste de Hiperparâmetros

O processo de treinamento foi realizado utilizando uma função de perda ponderada, devido ao desbalanceamento entre as classes positivas e negativas de sepse. Durante o treinamento, foram ajustados hiperparâmetros como a taxa de aprendizado, que controla a velocidade de ajuste do modelo, e o número de camadas, que define a profundidade da rede neural. Para evitar o sobreajuste, foram aplicados a técnica de *dropout*, que desativa aleatoriamente alguns neurônios durante o treinamento, forçando o modelo a aprender padrões mais gerais, e um mecanismo de parada antecipada, que interrompe o treinamento quando não há mais melhorias no desempenho (Yin et al., 2024).

6. Avaliação de Desempenho

A avaliação do modelo foi realizada por meio de métricas amplamente adotadas para classificação, incluindo a AUC-ROC e a AUPRC, que mensuram a capacidade discriminativa do modelo (McDermott et al., 2024). O F1-Score foi utilizado para evidenciar o equilíbrio entre precisão e sensibilidade, oferecendo uma visão crítica sobre o desempenho geral do modelo em termos de acurácia e minimização de falsos positivos e negativos (McDermott et al., 2024). Para uma análise mais específica dos acertos e erros, a matriz de confusão foi empregada, permitindo uma inspeção da performance do modelo em cada classe (McDermott et al., 2024). Além das métricas numéricas, gráficos de histórico de perda e distribuição das probabilidades previstas foram gerados, facilitando a interpretação da evolução do treinamento e a identificação de padrões nas previsões (McDermott et al., 2024).

RESULTADOS E DISCUSSÃO

Ao analisá-lo, foi identificado um problema relevante: a proporção entre casos positivos e negativos era extremamente desequilibrada, com aproximadamente uma sequência positiva para cada doze negativas. Além disso, as poucas sequências positivas presentes eram muito menores do que as negativas, o que agravava ainda mais o desbalanceamento. Apesar disso, esse tipo de desbalanceamento não é algo único deste conjunto, mas sim resultante da própria natureza dos dados, em que casos de sepse são relativamente raros quando comparados ao total de pacientes monitorados. Um dos principais obstáculos na etapa de preparação dos dados foi o alto índice de valores ausentes, especialmente em variáveis laboratoriais, como exames de sangue. Esses exames nem sempre eram coletados para todos os pacientes, o que resultava em colunas com longos trechos sem dados, o que, embora aplicando técnicas de interpolação tenham sido aplicadas para preencher lacunas, não eram capazes de reconstruir informações completamente ausentes.

FIGURA 1. Estatísticas do comprimento das sequências.

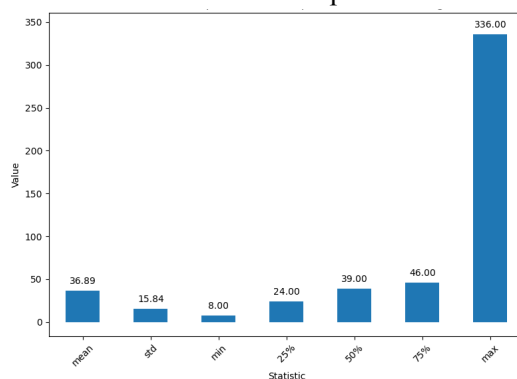
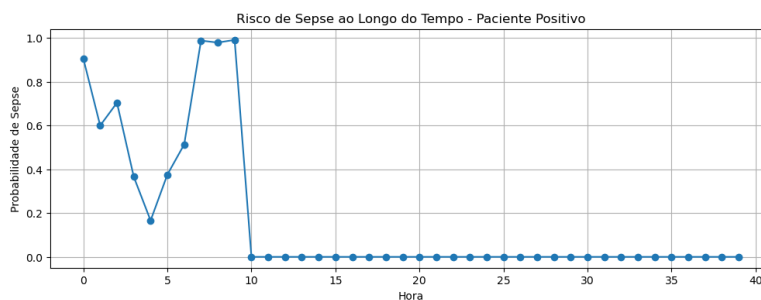


FIGURA 2. Simulação do sistema e exibição do risco de sepse por hora.



Na padronização das sequências temporais, optou-se por um comprimento fixo de 36 horas, por ser a média do tamanho das sequências, como aponta a Figura 1. No entanto, ao se aplicar esse tamanho fixo, aproximadamente mais da metade das horas de várias sequências positivas acabavam preenchidas com valores zerados, o que introduziu elementos artificiais nos dados. Esses valores de preenchimento, apesar de necessários para uniformizar as entradas, mais tarde se mostraram prejudiciais ao aprendizado da rede, especialmente quando foi tentado simular o comportamento do sistema em tempo real, como aponta a Figura 2.

A arquitetura inicial escolhida foi uma TCN (*Temporal Convolutional Network*), devido a sua capacidade especializada em lidar com dados sequenciais. Essa escolha foi motivada pela capacidade da TCN de capturar padrões temporais e evoluções numéricas ao longo do tempo, algo essencial para a detecção precoce da sepse. A TCN também incorpora conexões residuais e convoluções dilatadas, que expandem a recepção temporal sem a necessidade de redes excessivamente profundas. Apesar disso, outras arquiteturas que lidam com dados sequenciais, como LSTM (*Long Short-Term Memory*) ou GRU (*Gated Recurrent Unit*), são eficazes na modelagem de dependências temporais, o que as torna

boas opções para prever padrões que dependem do contexto passado, devido à sua capacidade de lidar com longas dependências temporais (Xiao et al., 2024), serão implementadas para comparar o desempenho entre diferentes abordagens.

TABELA 1. Métricas do modelo no conjunto de testes.

	Precisão	Sensibilidade	F1-Score	Suporte
0	1.0000	0.9795	0.9897	9972
1	0.7682	1.0000	0.8689	676
acurácia			0.9808	10648
média macro	0.8841	0.9898	0.9293	10648
média ponderada	0.9853	0.9808	0.9820	10648
AUROC: 0.9969				
AUPRC: 0.9387				

TABELA 2. Métricas do modelo no conjunto de validação.

	Precisão	Sensibilidade	F1-Score	Suporte
0	1.0000	0.9831	0.9915	5974
1	0.8043	1.0000	0.8915	676
acurácia			0.9842	6389
média macro	0.9021	0.9915	0.9415	6389
média ponderada	0.9873	0.9842	0.9850	6389
AUROC: 0.9974				
AUPRC: 0.9544				

Os resultados obtidos após o treinamento do modelo indicaram um bom desempenho, o que levanta a possibilidade de que possam estar inflados por algum viés ou problema: As métricas mostraram valores elevados em teste e validação, evidenciando consistência. A alta precisão para a classe negativa indica que poucos exemplos negativos foram classificados incorretamente, enquanto a sensibilidade elevada para a classe positiva mostra que quase todos os casos reais foram identificados. A acurácia geral reflete a proporção de acertos, e o F1-Score evidencia o equilíbrio entre precisão e sensibilidade. Métricas como AUROC e AUPRC demonstram excelente capacidade de distinção entre classes, mesmo com desequilíbrio dos dados (Tabelas 1 e 2). Esses resultados sugerem que o modelo aprendeu os padrões associados à condição. Para prevenir sobreajuste, foi implementada a interrupção precoce do treinamento quando não houve melhora significativa na perda de teste ou validação, evitando ciclos computacionais desnecessários. O balanceamento da função de perda, que penaliza mais erros na classe minoritária, teve efeito difícil de mensurar, mas é considerado boa prática em contextos com classes desbalanceadas.

A etapa de simulação em tempo real, que consiste em aplicar o modelo a pacientes não vistos durante o treinamento, revelou uma situação crítica. Esperava-se que o modelo fosse capaz de acompanhar, hora a hora, a progressão clínica do paciente e elevar a probabilidade de sepse quando sinais críticos comesçassem a aparecer. No entanto, observou-se que, em alguns casos, a partir da décima hora, a predição do modelo despencava para zero, mesmo em pacientes com diagnóstico confirmado de sepse, como mostra a Figura 2. Essa falha foi atribuída diretamente à distorção provocada pelos valores de *padding* com zeros, que não representavam dados reais, mas ainda assim eram interpretados como válidos pela rede, levando a predições artificiais e enganosas.

CONCLUSÕES

Os resultados deste estudo evidenciam que o uso de redes neurais artificiais constitui uma abordagem promissora para o diagnóstico precoce da sepse, alcançando métricas elevadas de desempenho que demonstram a capacidade do modelo em identificar padrões associados à evolução

clínica da condição. Contudo, ao confrontar tais resultados com os objetivos propostos, observa-se que, embora tenha havido êxito na etapa inicial de validação do modelo, desafios ainda persistem relacionados à dessemelhança dos dados clínicos, ao desbalanceamento entre classes e à presença de variáveis com ausência sistemática de registros. Esses fatores impactam a qualidade do modelo quando aplicado em cenários que simulam uma realidade hospitalar, limitando sua generalização. Assim, torna-se necessário avançar em estratégias mais sofisticadas de tratamento de dados faltantes, explorar diferentes arquiteturas de aprendizado profundo, além de comparar o desempenho com métodos tradicionais de aprendizado de máquina. Dessa forma, o trabalho reafirma a viabilidade da inteligência artificial como uma ferramenta de apoio à decisão clínica na sepse, ao mesmo tempo em que aponta a necessidade de uma abordagem metodológica mais abrangente para que sua aplicação prática se torne efetivamente confiável e integrada ao contexto hospitalar.

CONTRIBUIÇÕES DOS AUTORES

Autor 1: Curadoria, conceitualização, coleta e análise de dados, pesquisa, metodologia, desenvolvimento do modelo, redação do manuscrito original.

Autores 2 e 3: Orientação, revisão e contribuições intelectuais.

AGRADECIMENTOS

Agradeço à Coordenação de Pesquisa e Inovação do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Jacareí pelo apoio por meio da bolsa do PIBIFSP.

REFERÊNCIAS

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. arXiv preprint arXiv:1803.01271.

BOJER, Casper Solheim; MELDGAARD, Jens Peder. Kaggle forecasting competitions: An overlooked learning opportunity. *International Journal of Forecasting*, v. 37, n. 2, p. 587-603, 2021. DOI: <https://doi.org/10.1016/j.ijforecast.2020.07.007>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0169207020301114>. Acesso em: 17 set. 2025.

GOJAK, DŽ.; GVOŽĐAR, K.; HEĆIMOVIĆ, Z.; SMAJOVIĆ, A.; BEČIĆ, E.; DEUMIĆ, A.; SPAHIĆ BEĆIROVIĆ, L.; GURBETA POKVIĆ, L.; BADNJEVIĆ, A. The use of artificial intelligence in the diagnosis of neonatal sepsis. *IFAC-PapersOnLine*, v. 55, n. 4, p. 62–67, 2022. Disponível em: <https://doi.org/10.1016/j.ifacol.2022.06.010>. Acesso em: 28 jun. de 2025.

Johnson, A.; Pollard, T.; Mark, R. Banco de Dados Clínico MIMIC-III. Versão 1.4. *PhysioNet*, 2016. RRID:SCR_007345. Disponível em: <https://doi.org/10.13026/C2XW26>. Acesso em: 24 ago. 2025.

LEA, Colin; FLYNN, Michael D.; VIDAL, René; REITER, Austin; HAGER, Gregory D. Temporal Convolutional Networks for Action Segmentation and Detection. arXiv, 16 nov. 2016. Disponível em: <https://arxiv.org/abs/1611.05267>. Acesso em: 23 ago. 2025.

LIPTON, Zachary C.; KALE, David C.; WETZEL, Randall. Modeling Missing Data in Clinical Time Series with RNNs. 2016. Disponível em: <https://arxiv.org/abs/1606.04130> . Acesso em: 24 ago. 2025.

MANUAL MSD. Sepsis: causas, sintomas e tratamento. *Manual MSD para profissionais da saúde*, 2024. Disponível em: <https://www.msdmanuals.com/>. Acesso em: 28 set. 2024.

McDERMOTT, Matthew B. A.; ZHANG, Haoran; HANSEN, Lasse Hyldig; ANGELOTTI, Giovanni; GALLIFANT, Jack. A closer look at AUROC and AUPRC under class imbalance. 2024. Disponível em: <https://arxiv.org/abs/2401.06091>. Acesso em: 23 ago. 2025.

MOURA, J. M.; BERTOLLI, E. S.; PEREIRA, R. M.; FRUTUOSO, I. S.; WERNECK, A. L.; CONTRIN, L. M. Diagnóstico de sepse em pacientes após internação em unidade de terapia intensiva. *Arquivos de Ciência da Saúde*, v. 24, n. 3, p. 55–60, 2017.

PAIS, S. L.; AHMED, K. F.; SHARANYA; SHRIHASTHA; VARSHITH. Disease prediction using machine learning algorithms. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, v. 3, n. 1, p. 5–12, 2023.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Advances in Neural Information Processing Systems*, v. 32, 2019. Disponível em: <https://pytorch.org/>. Acesso em: 24 ago. 2025.

PEREIRA JUNIOR, G. A.; MARSON, F.; ABEID, M.; OSTINI, F. M.; SOUZA, S. H.; BASILE-FILHO, A. Fisiopatologia da sepse e suas implicações terapêuticas. *Medicina*, Ribeirão Preto, v. 31, n. 3, p. 349–362, 1998.

QUEIROZ, V. N. F. Abordagem e descrição metodológica do uso do prontuário eletrônico como forma de coleta de dados em pesquisa clínica em centro participante de estudo internacional e multicêntrico. *Einstein (São Paulo)*, v. 17, n. 3, p. eAO5543, 2019. Disponível em: https://doi.org/10.31744/einstein_journal/2019AE4791. Acesso em: 23 ago. 2025.

SHARAN, R.; SHARMA, D.; DHANWANTRI, A.; TANDON, S. Analysis of sepsis cases using deep learning. *Journal of Biomedical Informatics*, v. 14, n. 2, p. 67–78, 2023.

SHERSTINSKY, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, v. 404, 2020, p. 132306. ISSN 0167-2789. Disponível em: <https://doi.org/10.1016/j.physd.2019.132306>. Acesso em: 28 jun. de 2025.

SILVA, C. F. M.; MENDES, R. V.; COSTA, P. R. Diagnóstico precoce de sepse: uma análise da literatura. *Saúde e Pesquisa*, v. 15, n. 2, p. 123–135, 2022.

XIAO, Jue; BI, Shuochen; DENG, Tingting. Comparative Analysis of LSTM, GRU, and Transformer Models for Stock Price Prediction. 2024. Disponível em: <https://doi.org/10.48550/arXiv.2411.05790>. Acesso em: 17 set. 2025.

YIN, Changchang; CHEN, Pin-Yu; YAO, Bingsheng; WANG, Dakuo; CATERINO, Jeffrey; ZHANG, Ping. SepsisLab: early sepsis prediction with uncertainty quantification and active sensing. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '24)*, New York, NY, USA, 2024. p. 6158–6168. Disponível em: <https://doi.org/10.1145/3637528.3671586>. Acesso em: 28 jun. de 2025.