

**UNIVERSIDADE NOVE DE JULHO**  
**PROGRAMA DE PÓS-GRADUAÇÃO DE MESTRADO E DOUTORADO EM**  
**INFORMÁTICA E GESTÃO DO CONHECIMENTO**

**REINALDO RIBEIRO DE OLIVEIRA**

**NOVO ÍNDICE CLÍNICO-LABORATORIAL PARA PREDIÇÃO DE**  
**DETERIORAÇÃO EM ADULTOS HOSPITALIZADOS COM MODELO**  
**HÍBRIDO DE APRENDIZAGEM PROFUNDA**

**São Paulo**

**2026**

**UNIVERSIDADE NOVE DE JULHO**  
**PROGRAMA DE PÓS-GRADUAÇÃO DE MESTRADO E DOUTORADO EM**  
**INFORMÁTICA E GESTÃO DO CONHECIMENTO**

**REINALDO RIBEIRO DE OLIVEIRA**

**NOVO ÍNDICE CLÍNICO-LABORATORIAL PARA PREDIÇÃO DE**  
**DETERIORAÇÃO EM ADULTOS HOSPITALIZADOS COM MODELO**  
**HÍBRIDO DE APRENDIZAGEM PROFUNDA**

Tese de Doutorado, apresentada ao Programa de Pós-graduação em Informática e Gestão do Conhecimento da Universidade Nove de Julho, requisito obrigatório para obtenção do título de Doutor em Informática e Gestão do Conhecimento.

Orientador: Prof. Dr. Cleber Gustavo Dias

Linha de Pesquisa 2: Sistemas Inteligentes

**São Paulo**

**2026**

Oliveira, Reinaldo Ribeiro de.

Novo índice clínico-laboratorial para predição de deterioração em adultos hospitalizados com modelo híbrido de aprendizagem profunda. / Reinaldo Ribeiro de Oliveira. 2026.

177 f.

Tese (Doutorado) – Universidade Nove de Julho – UNINOVE, São Paulo, 2026.

Orientador (a): Prof. Dr. Cleber Gustavo Dias.

1. Índice de deteriorações clínicas. 2. Mortalidade. 3. Registros eletrônicos de saúde. 4. Aprendizado de máquina. 5. Séries temporais. 6. Inteligência artificial.

I. Dias, Cleber Gustavo.

II. Título

CDU 004

**ATA DE DEFESA DE TESE**

Ao décimo dia do mês de março de dois mil e vinte e seis, às 14h30, do programa de Pós-graduação, desta Universidade, reuniu-se em sessão pública a Comissão Julgadora da Tese de Doutorado de Reinaldo Ribeiro de Oliveira sob o título NOVO ÍNDICE CLÍNICO-LABORATORIAL PARA PREDIÇÃO DE DETERIORAÇÃO EM ADULTOS HOSPITALIZADOS COM MODELO HÍBRIDO DE APRENDIZAGEM PROFUNDA.

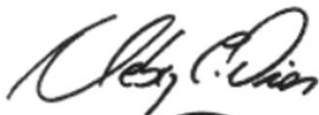
Integraram a comissão os professores: Prof. Dr. Cleber Gustavo Dias (UNINOVE), o Prof. Dr. Renato José Sassi - PPGI - UNINOVE, a Profa. Dra. Maria Elisabete Salvador Graziosi – UNIFESP e a Profa. Dra. Adriana Paula Jordão Isabella – UNINOVE, sob a presidência do primeiro, orientador da tese. A banca examinadora, tendo decidido aceitar a tese, passou à arguição pública do candidato. Encerrados os trabalhos, os examinadores deram parecer final sobre a tese.

	<b>Parecer</b>
Prof. Dr. Cleber Gustavo Dias	Aprovado
Prof. Dr. Renato José Sassi	Aprovado
Prof. Dr. Maria Elisabete Salvador Graziosi	Aprovado
Prof. Dr. Adriana Paula Jordão Isabella	Aprovado

**Parecer: O trabalho foi aprovado e a banca ainda atribuiu menção honrosa para o trabalho.**

Em conclusão, o candidato foi considerado aprovado, no grau de Doutor em Informática e Gestão do Conhecimento. E, para constar, eu, Prof. Dr. André Felipe Henriques Librantz, diretor do Programa de Mestrado e Doutorado em Informática e Gestão do Conhecimento, lavrei a presente ata que assino juntamente com os membros da banca examinadora.

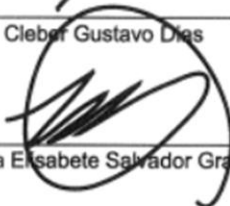
terça-feira, 10 de março de 2026



Prof. Dr. Cleber Gustavo Dias



Prof. Dr. Renato José Sassi



Prof. Dr. Maria Elisabete Salvador Graziosi



Prof. Dr. Adriana Paula Jordão Isabella



André Felipe Henriques Librantz

## RESUMO

**INTRODUÇÃO:** A deterioração clínica de pacientes, que podem levar a riscos de mortalidade, em especial nas unidades de terapia intensiva, está relacionada a um processo dinâmico de agravamento fisiológico com progressão de disfunção orgânica o qual pode ser monitorado ao longo do tempo por grandezas, ou indicadores, que podem ser obtidos dos prontuários clínicos dos pacientes. A avaliação da deterioração consiste em uma tarefa onerosa dada o comportamento clínico complexo e variável de cada paciente. O crescente volume de dados clínicos nos Prontuários Eletrônicos dos Pacientes (PEP), se caracterizam como registros longitudinais de informações de saúde dos pacientes. Com estes dados são possíveis utilizar modelos de protocolos de deteriorações clínicas, que podem observar e acompanhar a evolução fisiológica em pacientes hospitalizados. Com uso do prontuário eletrônico do paciente, os registros eletrônicos, podem gerar *insights* e previsões de eventos futuros. Este estudo, justifica-se pela necessidade de desenvolvimento de um novo índice fundamentado em inteligência artificial híbrida, capaz de processar dados clínicos heterogêneos temporais de registros eletrônicos de saúde. **OBJETIVO:** Desenvolver e testar um índice composto por dados clínicos e de laboratório para a previsão de deterioração em pacientes adultos hospitalizados com o uso de algoritmos de aprendizagem de máquina. **MÉTODO:** O estudo é de natureza tecnológica experimental exploratório. Para a condução dos experimentos foram selecionados 1.100.000 pacientes a partir de 2 *datasets* extraídos inicialmente da base de dados MIMIC-III. O estudo foi organizado em três etapas: Na primeira etapa foi realizada uma revisão de escopo, na segunda etapa foi determinado o índice de deterioração do paciente adulto a partir dos dados clínicos e de laboratório com séries temporais mais relevantes tratados estatisticamente e na última etapa foram realizados os experimentos computacionais, usando uma abordagem híbrida de redes neurais convolucionais e redes de memória de curto e longo prazo, a fim de prever o índice de deterioração em passos de tempo futuros. Construiu-se então os modelos matemáticos necessários para cálculo dos pesos das variáveis que fazem parte do índice de deterioração. **RESULTADOS:** O modelo de redes neurais híbridas, recebeu ajustes em 10 cenários diferentes para avaliação de seus desempenhos, com ajustes de hiper parâmetros em “*Seed, Past\_H, Future\_H, Val\_Split, Epochs e Batch Size*”, onde foi possível demonstrar a previsão de eventos futuros com uma série temporal retroativa de registros e treinamentos para prever a sexta (6) hora, décima segunda (12) hora, décima oitava (18) hora e vigésima quarta (24) hora. O desempenho obteve métricas globais de MAE = 9,36 - 9,89 e MAPE = 24,19 - 25,96%. **CONCLUSÕES:** Os desempenhos do modelo híbrido, apresentou resultados muito próximos com hiper parâmetros semelhantes, com pacientes de bases diferentes, como MIMIC-III e MIMIC-IV. Os melhores resultados foram observados na previsão das próximas 6 horas futuras. Por fim, sugerimos que a pesquisa tem potencial de ineditismo explorado na condução do trabalho, que projeta contribuições científicas para a comunidade acadêmica, para a sociedade e para os processos assistenciais, e que os resultados podem antecipar eventos clínicos e oferece oportunidades de intervenções e condutas no ambiente hospitalar para a segurança e proteção dos pacientes.

**Palavras-chaves:** Índice de Deteriorações Clínicas, Mortalidade, Registros Eletrônicos de Saúde, Aprendizado de Máquina, Séries Temporais, Inteligência Artificial.

## ABSTRACT

**INTRODUCTION:** Clinical deterioration in patients, which may lead to increased mortality risk, particularly in intensive care units, is associated with a dynamic process of physiological decline characterized by the progression of organ dysfunction. This process can be monitored over time through measurable indicators obtained from patients' clinical records. Assessing clinical deterioration is a demanding task due to the complex and variable clinical behavior of each patient. The growing volume of clinical data within Electronic Health Records (EHRs) is characterized as longitudinal patient health information. These data enable the application of clinical deterioration protocols capable of observing and tracking physiological progression in hospitalized patients. The use of EHRs allows electronic records to generate insights and predictions of future clinical events. This study is justified by the need to develop a novel index grounded in hybrid artificial intelligence capable of processing heterogeneous temporal clinical data derived from electronic health records. **OBJECTIVE:** To develop and test an index composed of clinical and laboratory data for predicting deterioration in hospitalized adult patients using machine learning algorithms. **METHODS:** This study has an exploratory experimental technological design. A total of 1,100,000 patients were selected from two datasets originally extracted from the MIMIC-III database. The study was structured into three stages: (i) a scoping review, (ii) development of an adult patient deterioration index based on the most relevant clinical and laboratory time-series data treated statistically, and (iii) computational experiments using a hybrid approach combining convolutional neural networks (CNNs) and long short-term memory (LSTM) networks to predict the deterioration index at future time steps. Mathematical models were constructed to calculate the weights of variables composing the deterioration index. **RESULTS:** The hybrid neural network model was evaluated across 10 different scenarios with hyperparameter tuning (Seed, Past\_H, Future\_H, Val\_Split, Epochs, and Batch\_Size). The model demonstrated the ability to predict future events using retrospective time-series records to forecast 6-hour, 12-hour, 18-hour, and 24-hour horizons. Global performance metrics achieved were MAE ranging from 9.36 to 9.89 and MAPE from 24.19% to 25.96%. **CONCLUSIONS:** The hybrid model showed consistent performance under similar hyperparameters across different patient cohorts from MIMIC-III and MIMIC-IV. The best results were observed for 6-hour-ahead predictions. The study presents innovative potential and offers scientific contributions to the academic community, society, and healthcare delivery processes. The findings suggest that the proposed approach can anticipate clinical events and provide opportunities for timely interventions and clinical decision-making in hospital settings, thereby enhancing patient safety.

**Keywords:** Clinical Deterioration Index; Mortality; Electronic Health Records; Machine Learning; Time Series; Artificial Intelligence.

Dedico este trabalho à minha esposa Simone, a minha mãe Suely, ao meu irmão Rodiney e ao meu orientador o Prof. Dr. Cleber Gustavo Dias. Vocês representam minhas motivações centrais para o desenvolvimento e conclusão deste trabalho.

## AGRADECIMENTOS

Agradeço, em primeiro lugar, a Deus, meus mentores espirituais, que sempre estiveram ao meu lado, orientando e iluminando meus passos e inspirações, colocando as pessoas e oportunidades certas em meu caminho.

A Universidade Nove de Julho (UNINOVE) desde a graduação, pelo apoio e pela oportunidade de crescimento e aprimoramento acadêmico, pessoal e profissional, assim como pela bolsa de estudos.

A minha esposa e companheira, Simone Rodrigues de Oliveira, que de maneira compreensiva, sempre me apoiou, incentivou muito, para a produção de conteúdos e cumprimento de prazos e metas.

Aos meus pais, Suely Alves de Oliveira por estar comigo e Ildebrando Ribeiro de Oliveira (*in memoriam*), minha base ideológica, por quem sempre me inspirei para chegar até aqui. A minha sogra Silvana Rodrigues por ser minha fiel apoiadora;

A empresa SPDM (Associação Paulista para o Desenvolvimento da Medicina) minhas lideranças diretas como a Profa Dra Elizabeth Akemi Nishio, Msc. Vanessa Ceravolo Gurgel e Msc. Cristiane Aparecida Betta e aos demais colegas de Núcleo Corporativo de Enfermagem. Aos meus demais colegas diretores, gerentes e supervisores de enfermagem de toda a corporação, que me apoiaram alcançar meus objetivos.

Ao meu inspirador amigo, um ser de luz, que sempre vai brilhar em minha vida, Toddy de Iervolino. Você transformou minha vida e tem eternamente seu espaço reservado em minhas lembranças, por onde estiver.

As professores do Programa de Pós Graduação (PPGI), suas contribuições foram fundamentais para a construção de uma pesquisa fora do meu habitat das Ciências da Saúde. Muito honrado, por ter sido possível e necessário desbravar caminhos desconhecidos nesta trajetória.

Em especial, cumprimento e agradeço ao Prof Dr Renato Jose Sassi, Profa Dra Maria Elisabete Salvador Graziosi e Profa Dra Adriana de Paula Jordão Isabella, além de estarem nas bancas de avaliações da pesquisa, vocês são minhas inspirações de vida por onde decidi seguir como pesquisador.

Ao Prof Dr Cleber Gustavo Dias, meu mentor e orientador, pelo apoio, suporte, confiança, paciência e por acreditar em mim. Gratidão para toda a vida.

Por fim, os meus sinceros agradecimentos a todos que de alguma forma, contribuíram coma minha jornada acadêmica.

"O correr da vida embrulha tudo. A vida é assim: esquenta e esfria, aperta e daí afrouxa, sossega e depois desinquieta. O que ela quer da gente é coragem."

**Guimarães Rosa** (em *Grande Sertão: Veredas*)

## FIGURAS

01	DETERIORAÇÃO DE UM PACIENTE AO LONGO DE 8 HORAS CULMINANDO EM PARADA CARDÍACA E MORTE	42
02	ESTRUTURA DE UM NEURÔNIO BIOLÓGICO	43
03	ESTRUTURA DE UM NEURÔNIO ARTIFICIAL	44
04	EXEMPLOS DA ESTRUTURA DE UMA REDE NEURAL CONVOLUCIONAL (CNN)	50
05	ESTRUTURA DE UMA REDE NEURAL RECORRENTE	51
06	ARQUITETURA DE UMA REDE NEURAL LSTM	52
07	ESTRUTURA DE UMA REDE NEURAL LSTM	53
08	CONEXÕES DE UMA REDE NEURAL LSTM	54
09	ORGANIZAÇÃO DAS FASES DA REVISÃO DE ESCOPO DE PESQUISA DA LITERATURA	56
10	CHECKLIST DE REVISÃO DE ESCOPO BASEADO EM MÉTODO PRISMA	59
11	TRABALHOS DERIVADOS DE: SCALABLE AND ACCURATE DEEP LEARNING WITH ELECTRONIC HEALTH RECORDS	70
12	TRABALHOS DERIVADOS DE: THE VALUE OF VITAL SIGN TRENDS IN PREDICTING AND MONITORING CLINICAL DETERIORATION: A SYSTEMATIC REVIEW.	73
13	TRABALHOS DERIVADOS DE: DEEP INTERPRETABLE EARLY WARNING SYSTEM FOR THE DETECTION OF CLINICAL DETERIORATION.	75
14	TRABALHOS DERIVADOS DE: THE PROMISE OF ARTIFICIAL INTELLIGENCE: TRABALHOS DERIVADOS: THE PROMISE OF ARTIFICIAL INTELLIGENCE: A REVIEW OF THE OPPORTUNITIES AND CHALLENGES OF ARTIFICIAL INTELLIGENCE IN HEALTHCARE	77
15	TRABALHOS DERIVADOS DE: MIMIC-III, A FREELY ACCESSIBLE CRITICAL CARE DATABASE	79
16	TRABALHOS DERIVADOS DE: BIOBERT: A PRE-TRAINED BIOMEDICAL LANGUAGE REPRESENTATION MODEL FOR BIOMEDICAL TEXT MINING	81
17	CARACTERIZAÇÃO METODOLÓGICA DA PESQUISA	93
18	EXIBIÇÃO DO PLANEJAMENTO DAS FASES NA METODOLOGIA UTILIZADA NA PESQUISA	95
19	METODOLOGIA KDD (KNOWLEDGE DISCOVERY IN DATABASE)	101

20	REDE ORIGINÁRIA DE PROCESSAMENTO DOS DADOS DO ÍNDICE PROPOSTO	103
21	DIAGRAMA DE EXPOSIÇÃO DA ORIGEM DOS DADOS E PROCESSAMENTO DO ÍNDICE	105
22	REPRESENTAÇÃO VISUAL DA JUNÇÃO DE ICLÍNICO + ILABORATÓRIO = IPED	106
23	REPRESENTAÇÃO VISUAL DA JUNÇÃO DE ICLÍNICO + ILABORATÓRIO + ITEXTO = IPET	106
24	VARIÁVEIS DOS SINAIS VITAIS E CÁLCULO DO ICLÍNICO	109
25	VARIÁVEIS DOS EXAMES DE LABORATÓRIO E CÁLCULO DE ILABORATÓRIO	110
26	VARIÁVEIS DOS RESULTADOS DE ANÁLISE DOS TEXTOS, NÃO ESTRUTURADOS COM ANOTAÇÕES E EVOLUÇÕES DE ENFERMEIROS E MÉDICOS	111
27	ARQUITETURA DE UMA REDE NEURAL HÍBRIDA COM USO DE LSTM - CNNS	113
28	RESULTADOS DOS ACHADOS DE PESO DAS VARIÁVEIS NO ICLÍNICO	118
29	RESULTADOS DOS ACHADOS DE PESO DAS VARIÁVEIS NO ILABORATÓRIO	119
30	MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED - MÉDIA	120
31	MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED - PONDERADO	121
32	MÉTRICAS IDENTIFICADAS NO ACUMULADO DE PESO DO CÁLCULO DE IPET	122
33	PROJEÇÃO ILUSTRATIVA DA PROPOSTA DE PESQUISA	123
34	EXIBIÇÃO AMOSTRAL DO DATASET, GERADO A PARTIR DOS DADOS REAIS COM OS CÁLCULOS DE IPED	124
35	TESTE DE AVALIAÇÃO DOS MELHORES PARÂMETROS PARA A REDE CNN-LSTM	125
36	PREVISÃO: USANDO COMO ALVO DO TREINAMENTO E O RESÍDUO (DELTA) PREDIÇÃO T+1 (6 HORAS)	127
37	PREVISÃO: USANDO ALVO DO TREINAMENTO E O RESÍDUO (DELTA) PREDIÇÃO T + 2 (12 HORAS)	127
38	PREVISÃO: USANDO ALVO DO TREINAMENTO E O RESÍDUO (DELTA) PREDIÇÃO T + 3 (18 HORAS)	131

39	PREVISÃO: USANDO ALVO DO TREINAMENTO E O RESÍDUO (DELTA) PREDIÇÃO T + 4 (24 HORAS)	133
40	PREVISÃO: USANDO COMO ALVO DO TREINAMENTO SEM RESÍDUO, PREDIÇÃO T + 1 (6 HORAS)	135
41	PREVISÃO: USANDO COMO ALVO DO TREINAMENTO SEM RESÍDUO, PREDIÇÃO T + 2 (12 HORAS)	137
42	PREVISÃO: USANDO COMO ALVO DO TREINAMENTO SEM RESÍDUO, PREDIÇÃO T + 3 (18 HORAS)	139
43	PREVISÃO: USANDO COMO ALVO DO TREINAMENTO SEM RESÍDUO, PREDIÇÃO T + 4 (24 HORAS)	141
44	PREVISÃO: COMO ALVO DO TREINAMENTO COM RESÍDUO DELTA PREDIÇÃO T + 1 (6 HORAS)	143
45	PREVISÃO: COMO ALVO DO TREINAMENTO COM RESÍDUO DELTA PREDIÇÃO T + 2 (12 HORAS)	145
46	COEFICIENTE DE REGRESSÃO PARA O MELHOR CASO ( t + 1 )	147
47	PREDIÇÃO DA DETERIORAÇÃO DO PACIENTE (VARIAÇÃO DO IPED) PARA 5 PACIENTES DA BASE MIMIC-IV	149

## QUADROS

01	PUBLICAÇÕES QUE ABORDARAM: APRENDIZAGEM DE MÁQUINA EM REGISTROS ELETRÔNICOS DE SAÚDE	67
02	PUBLICAÇÕES QUE ABORDARAM: DETERIORAÇÕES CLÍNICAS OBSERVADAS POR APRENDIZAGEM DE MÁQUINA	71
03	PUBLICAÇÕES QUE ABORDARAM: DETERIORAÇÕES CLÍNICAS OBSERVADAS POR APRENDIZAGEM DE MÁQUINA	74
04	PUBLICAÇÕES QUE ABORDARAM: INTELIGÊNCIA ARTIFICIAL PARA PREVER DETERIORAÇÕES CLÍNICAS	76
05	PUBLICAÇÕES QUE ABORDARAM: BANCO DE DADOS UTILIZADOS EM SAÚDE	78
06	PUBLICAÇÕES QUE ABORDARAM: ENTIDADES NOMEADAS COM MODELOS PRÉ TREINADOS	79
07	PUBLICAÇÕES QUE ABORDARAM: INSTRUMENTOS DE AVALIAÇÕES DE DETERIORAÇÕES CLÍNICAS	82
08	PUBLICAÇÕES QUE ABORDARAM: PREDIÇÕES DE EVENTOS COM APRENDIZAGEM DE MÁQUINA	83
09	PUBLICAÇÕES QUE ABORDARAM: REDES NEURAIS USADAS EM REGISTROS ELETRÔNICOS DE SAÚDE	85
10	PUBLICAÇÕES QUE ABORDARAM: PROCESSAMENTO DE LINGUAGEM NATURAL UTILIZADAS EM REGISTROS ELETRÔNICOS DE SAÚDE	86
11	PUBLICAÇÕES QUE ABORDARAM: OUTRAS CLASSES RELACIONADAS AO APRENDIZAGEM DE MÁQUINA	88
12	PUBLICAÇÕES QUE ABORDARAM: REVISÕES SISTEMÁTICAS RELACIONADAS AO APRENDIZAGEM DE MÁQUINA	89
13	PUBLICAÇÕES QUE ABORDARAM: ESTUDOS COM MÉTODO PRISMA	90
14	EXIBIÇÃO DAS PUBLICAÇÕES PRODUZIDAS DURANTE A ELABORAÇÃO DA TESE	116

## TABELAS

01	PUBLICAÇÕES QUE NORTEARAM A PESQUISA	29
02	TABELA COM A DESCRIÇÃO DAS PALAVRAS CHAVES E MESH UTILIZADOS NA PESQUISA	98
03	PROPOSTA DE PARAMETRIZAÇÃO INICIAL PARA DE T + 1 (6 HORAS)	114
04	RESULTADOS DOS TESTES USANDO OS DADOS DO MIMIC-IV	147

## GRÁFICOS

01	TOTALIZAÇÃO DE BUSCAS EM BASES DE DADOS	60
02	TOTALIZAÇÃO DE BUSCAS EM BASES SELECIONADAS	60
03	TOTALIZAÇÃO DE PUBLICAÇÕES SELECIONADAS POR ANO	61
04	TOTALIZAÇÃO DE PUBLICAÇÕES MAPEADAS POR CLASSES	62
05	TOTALIZAÇÃO DE PUBLICAÇÕES MAPEADAS POR PAÍSES DE ORIGENS	63
06	RANKING DOS PERIÓDICOS MAIS BEM CLASSIFICADOS, SEGUNDO NÚMERO DE CITAÇÕES	65
07	GEO-MAPEAMENTO DAS PUBLICAÇÕES PAÍSES DE ORIGENS DOS ESTUDOS	66

## EQUAÇÕES

01	VETOR DE CÁLCULO DE PESOS DAS VARIÁVEIS	45
02	ATUALIZAÇÃO DE CÁLCULO DOS PESOS	45
03	MODELO MATEMÁTICO PARA CÁLCULO DAS VARIÁVEIS	105
04	MODELO MATEMÁTICO PARA CÁLCULO E DEFINIÇÃO DO PESO DE CADA VARIÁVEL	107
05	MODELO MATEMÁTICO PARA CÁLCULO DE RESULTADOS DE IPED	111
06	MODELO MATEMÁTICO PARA CÁLCULO DE RESULTADOS DE IPET	111

## LISTA DE ABREVIATURAS

AI	<i>Artificial Inteligency</i>
BIDMC	<i>Data warehouse seguro do Beth Israel Deaconess Medical Center</i>
BILLSTM	<i>Rede Neural Long Short-Term Memory Bidirecional</i>
BIOBERT	<i>Modelos de entidades nomeadas pré-treinados usados em textos da saúde</i>
BRADEN	<i>Instrumento de avaliação para determinar o risco de lesões por pressão</i>
CINAHL	<i>The Cumulative Index to Nursing and Allied Health Literature</i>
CLINICALBERT	<i>Modelos de entidades nomeadas pré-treinados usados em textos da saúde</i>
CNNs	<i>Convolutional neural network</i>
DEEPHOSP	<i>Índice de Predição de Deteriorações Clínicas no Paciente Adulto</i>
DL	<i>Deep Learning</i>
EHRs	<i>Eletronic Health Records</i>
EMBASE	<i>Excerpta Medica dataBASE</i>
FC	<i>Frequência Cardíaca</i>
FR	<i>Frequência Respiratória</i>
GLASGOW	<i>Score de Avaliação do Nível de Consciência</i>
HIPAA	<i>Health Insurance Portability and Accountability Act</i>
IA	<i>Inteligência Artificial</i>
ICD	<i>Classificação Internacional de Doenças</i>
IEEE Explore	<i>Um banco de dados de pesquisa relacionados em ciência da computação, engenharia elétrica e eletrônica</i>
KCH	<i>King's College Hospital</i>
KDD	<i>Knowledge Discovery in Databases</i>
LOINC	<i>Nomes e Códigos de Identificadores de Observação Lógica</i>
LSTM	<i>Rede Neural Long Short-Term Memory</i>
MEDLINE	<i>Medical Literature Analysis and Retrievel System online</i>
MEOWS	<i>Modified Early Obstetric Warning Score</i>
MEWS	<i>Modified Early Warning Score</i>
MIMIC-III	<i>US Medical Information mart for Intensive Care</i>
MIT	<i>Massachusetts Institute of Technology</i>
ML	<i>Machine Learning</i>
MORSE	<i>Instrumento de avaliação para determinar o risco de queda do paciente</i>
NER	<i>Named Entity Recognition</i>
NEWS	<i>National Early Warning Score</i>
PAD	<i>Pressão Arterial Diastólica</i>
PAS	<i>Pressão Arterial Sistólica</i>
PEP	<i>Prontuário Eletrônico do Paciente</i>
PEWS	<i>Pediatric Early Warning Score</i>
PLN	<i>Processamento de Linguagem Natural</i>
PNL	<i>Natural Language Processing</i>
PRISMA	<i>Preferred Reporting Items for Systematic reviews and Meta-Analyses</i>
PUBMED	<i>Um banco de dados ciências biológicas e biomédicas</i>
PUBMEDBERT	<i>Modelos de entidades nomeadas pré-treinados usados em textos da saúde</i>
PULSEDB	<i>Grande conjunto de dados limpos do MIMIC-III</i>
QSOFA	<i>Score for In-Hospital Mortality Among Adults</i>
REN	<i>Reconhecimento de Entidades Nomeadas</i>
RES	<i>Registros Eletrônicos de Saúde</i>
SAPS	<i>Simplified Acute Physiology Score</i>

SatO <sup>2</sup>	<i>Saturação de oxigênio no sangue</i>
SLAM	<i>South London and Maudsley</i>
SNOMED CT	<i>Nomenclatura Sistematizada de Medicina - Termos Clínicos</i>
SOFA	<i>Sequential Organ Failure Assessment</i>
TEMP	<i>Temperatura Corporal</i>
UMLS	<i>Sistema Unificado de Linguagem Médica</i>
VITALDB	<i>Software de registros eletrônicos de saúde</i>

## SUMÁRIO

<b>1.0</b>	<b>INTRODUÇÃO</b>	20
1.1	MOTIVAÇÕES DA PESQUISA	25
1.2	JUSTIFICATIVAS DA PESQUISA	27
1.3	IDENTIFICAÇÃO DAS LACUNAS DE PESQUISA	33
1.4	QUESTÃO DE PESQUISA	37
1.5	OBJETIVOS	37
	1.5.1 – OBJETIVO GERAL	37
	1.5.2 – OBJETIVOS ESPECÍFICOS	37
1.6	DELIMITAÇÃO DO TEMA DE PESQUISA	38
1.7	ORGANIZAÇÃO DO TRABALHO DE PESQUISA	38
<b>2.0</b>	<b>REVISÃO DA LITERATURA E FUNDAMENTAÇÃO TEÓRICA</b>	39
2.1	A DETERIORAÇÃO CLÍNICA DE PACIENTES EM UTI E O USO DE INDICADORES PARA A SUA AVALIAÇÃO	40
2.2	FUNDAMENTOS DAS REDES NEURAI ARTIFICIAIS	43
2.3	REGISTROS ELETRÔNICOS DE SAÚDE (RES)	45
2.4	MODELOS PREDITIVOS EM SAÚDE	46
2.5	DESENVOLVIMENTO TÉCNICO DOS ALGORÍTMOS	48
2.6	TÉCNICAS DE RECONHECIMENTO ENTIDADES NOMEADAS EM SAÚDE	48
2.7	REDES NEURAI CONVOLUCIONAIS	49
2.8	REDES NEURAI RECORRENTES	50
2.9	REDES NEURAI RECORRENTES LSTM	51
2.10	REDES NEURAI RECORRENTES LSTM BIDIRECIONAIS	53
2.11	REVISÃO DE ESCOPO DA LITERATURA SOBRE O USO DE INTELIGÊNCIA ARTIFICIAL EM REGISTROS ELETRÔNICOS DE SAÚDE	55
<b>3.0</b>	<b>MATERIAIS E MÉTODOS</b>	92
3.1	CARACTERIZAÇÃO DA NATUREZA DE PESQUISA	92
3.2	DESCRIÇÃO DAS PRINCIPAIS ETAPAS DA PESQUISA	94
	3.2.1 – DETALHAMENTO DA ETAPA 2.1 DA FASE 2	97
	3.2.2 – ESTRATÉGIAS DE PESQUISAS NA FASE DE FUNDAMENTAÇÃO	97
	3.2.3 – DETALHAMENTO DA ETAPA 2.2 DA FASE 2	99
3.3	LIMPEZA E TRATAMENTO DOS DADOS	101
	3.3.1 – ORIGEM DOS DADOS	103
3.4	DETERMINAÇÃO ÍNDICE DA PREDIÇÃO CLÍNICA DOS PACIENTES	106
3.5	PADRONIZAÇÃO DE CÁLCULO E EQUAÇÕES DOS MODELOS	107
3.6	DEFINIÇÃO DO PESO DE CADA VARIÁVEL	108
3.7	DEFINIÇÃO DO ÍNDICE ICLÍNICO	109
3.8	DEFINIÇÃO DO ÍNDICE ILABORATÓRIO	110
3.9	DEFINIÇÃO DO ÍNDICE ITEXTO	111
3.10	DEFINIÇÃO DOS ÍNDICES COMBINADOS DE: IPED OU IPET	111
3.11	DETALHAMENTO DA ETAPA 2.3 DA FASE 2	112
3.12	DEFINIÇÃO DA ARQUITETURA HÍBRIDA COM USO DE REDE NEURAL LSTM - CNNs	113
3.13	DEFINIÇÃO DOS HIPERPARÂMETROS DA REDE NEURAL CNN – LSTM PROPOSTO PARA OS EXPERIMENTOS	114
<b>4.0</b>	<b>RESULTADOS e DISCUSSÕES</b>	116
4.1	RESULTADOS PRELIMINARES DE PUBLICAÇÕES PRODUZIDAS NA PRIMEIRA FASE DA PESQUISA	116
4.2	RESULTADOS OBTIDOS NA SEGUNDA ETAPA DA PESQUISA	117

4.2.1	- RESULTADOS OBTIDOS PARA A DETERMINAÇÃO DO IPED A PARTIR DAS PARCELAS ICLINICO E ILABORATÓRIO	117
4.3	CATEGORIZAÇÃO DOS PESOS DOS COEFICIENTES DAS VARIÁVEIS DE ICLÍNICO	118
4.4	CATEGORIZAÇÃO DOS PESOS DOS COEFICIENTES DAS VARIÁVEIS DE ILABORATÓRIO	119
4.5	MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED - MÉDIA	120
4.6	MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED – PONDERADO	121
4.7	MÉTRICAS IDENTIFICADAS NO ACUMULADO DE PESO DO CÁLCULO DE IPET	122
4.8	RESULTADOS DA PREDIÇÃO DO IPED OBTIDOS COM OS ALGORITMOS CNN - LSTM	125
4.9	PROPOSTAS DOS CENÁRIOS TESTADOS NOS EXPERIMENTOS COMPUTACIONAIS	126
4.10	PROPOSTAS DOS CENÁRIOS TESTADOS - 1	127
4.12	PROPOSTAS DOS CENÁRIOS TESTADOS - 2	129
4.13	PROPOSTAS DOS CENÁRIOS TESTADOS - 3	131
4.14	PROPOSTAS DOS CENÁRIOS TESTADOS – 4	133
4.15	PROPOSTAS DOS CENÁRIOS TESTADOS – 5	135
4.16	PROPOSTAS DOS CENÁRIOS TESTADOS – 6	137
4.17	PROPOSTAS DOS CENÁRIOS TESTADOS – 7	139
4.18	PROPOSTAS DOS CENÁRIOS TESTADOS – 8	141
4.19	PROPOSTAS DOS CENÁRIOS TESTADOS – 9	143
4.20	PROPOSTAS DOS CENÁRIOS TESTADOS – 10	145
4.21	PROPOSTA DE TESTE DO MELHOR MODELO TREINADO COM OS DADOS DO MIMIC-III USANDO OS DADOS DO DATASET MIMIC-IV	147
4.22	CONSIDERAÇÕES FINAIS DOS RESULTADOS DOS EXPERIMENTOS COMPUTACIONAIS	150
4.23	PERSPECTIVAS PARA O FUTURO NOS AMBIENTES HOSPITALARES PERSPECTIVAS DA PROPOSTA DE PESQUISA COM A	151
4.24	INTEROPERABILIDADE DE SISTEMAS COM REGISTROS ELETRÔNICOS DE SAÚDE	152
4.25	O CONTEXTO BRASILEIRO QUE ABORDA O TEMA DE PESQUISA	154
4.26	LIMITAÇÕES DA PESQUISA	156
4.27	CONTRIBUIÇÕES DA PESQUISA (Pontos Fortes)	157
<b>5.0</b>	<b>CONCLUSÕES</b>	159
<b>6.0</b>	<b>ANEXOS</b>	161
<b>7.0</b>	<b>REFERÊNCIAS</b>	166

## 1.0 INTRODUÇÃO

No atendimento assistencial, em particular para os diferentes cuidados no âmbito da saúde, o foco está sempre orientado para a segurança do paciente e para a vigilância contínua das suas condições clínicas. Quando se trata da observação da chamada deterioração clínica de pacientes internados em unidades de terapia intensiva (UTI), por exemplo, tal condição não deve ser vista apenas pelo desfecho final (morte ou alta), mas principalmente pela sequência de complicações que o paciente apresenta ou pode indicar em um dado período (**VICENT et al. 1996**).

As manifestações clínicas dos pacientes podem aparecer, ora em resultados dos sinais vitais, ora em exames laboratoriais, ora em respostas orgânicas, como: dificuldades respiratórias, alterações de pressão arterial, perdas de consciência e confusões mentais ou até mesmo com eventos simultâneos (**JURAEV et al., 2022**).

A deterioração clínica laboratorial do paciente, segundo **PIRRACCHIO et al. (2019)** pode acontecer quando um paciente está com tendência de declínio do seu estado de estabilidade clínica, frente ao equilíbrio fisiopatológico no funcionamento do corpo humano, reflexão também destacada por **PADILLA et al (2018)**.

Na prática das atividades hospitalares, pode-se observar que as disfunções cardiológicas por exemplo, podem aparecer até 24 horas antes dos eventos consumados, segundo **CONSIDINE et al (2024)**, sabendo que, prever as deteriorações clínicas, pode ser fundamental para que as intervenções aconteçam em tempo hábil, a fim de amenizar ou até salvar a vida dos pacientes comprometidos.

Há muito tempo, e conforme destacado por **VICENT et al (1996)**, sabe-se que a deterioração clínica laboratorial deve ser entendida como um processo, e não como um evento único, de tal modo que possíveis disfunções orgânicas devem ser vistas ao longo de um tempo contínuo e não algo que está apenas “presente” ou “ausente”.

Em episódios de instabilidade, por exemplo, a equipe assistencial realiza a monitorização de eventos clínicos utilizando escalas e índices amplamente validados no ambiente hospitalar.

O estudo conduzido por **LU et al (2022)**, por exemplo, realizou um comparativo observacional de escalas, tradicionalmente conhecidas na assistência hospitalar, que apoiam a observação de eventos de deteriorações clínicas, que são relativamente conhecidas por profissionais de saúde, como o MEWS (*Modified Early Warning Score*) usado em pacientes clínicos adultos e o NEWS (*National Early Warning Score*) usado em pacientes clínicos adultos com tendência a complicações respiratórias.

A escala PEWS (*Pediatric Early Warning Score*) é usada em pacientes clínicos pediátricos, o MEOWS (*Modified Early Obstetric Warning Score*) é usado em gestantes, o SAPS (*Simplified Acute Physiology Score*) usado em pacientes em unidades de terapia intensiva adulto, o SOFA (*Sequential Organ Failure Assessment*) empregado em pacientes em unidades de terapia intensiva adulto.

A escala QSOFA (*Score for In-Hospital Mortality Among Adults*), por exemplo, usada em pacientes em unidades de terapia intensiva adulto, entre outras escalas, baseiam em sistemas simples de pontuação agregada, quando são associados a medidas fisiológicas já registradas nas práticas dos profissionais de saúde. (ASIIMWE et al., 2020, LUO et al., 2021 e KIA et al., 2020).

Cabe destacar que, seis principais parâmetros fisiológicos simples, formam a base dos sistemas de pontuação (score) a depender do tipo e perfil do paciente avaliado, são eles:

1. Frequência respiratória (FR);
2. Saturação de oxigênio no sangue (SatO<sup>2</sup>);
3. Pressão arterial sistólica e diastólica (PAS) e (PAD);
4. Frequência cardíaca (FC);
5. Nível de consciência (Glasgow);
6. Temperatura corporal (Temp);

Na prática clínica, quando um paciente apresenta confusão, desorientação e/ou agitação de início recente, considerando que seu estado mental era normal há horas, isso pode sugerir o início de uma sutil complicação, destacado por LEE et al. (2020).

Outro exemplo clássico, observado na prática assistencial, quando o paciente responde às perguntas de forma coerente, mas há algum quadro de confusão, desorientação e/ou agitação, pela escala de Coma de Glasgow, podemos atribuir as pontuações a cada parâmetro pelas respostas que são observadas, com a pontuação, refletindo o quão o parâmetro varia em relação ao normal. (BELL et al. 2021)

Atualmente, temos protocolos disponíveis e recursos manuais para prever deteriorações clínicas nos pacientes, como descrito por LU et al. (2022) e RUIZ et al., (2022), também, dispomos de recursos computacionais, ou ferramentas, a exemplo dos algoritmos de inteligência artificial, e acredita-se que seja possível também desenvolver sistemas inteligentes de apoio, que possam contribuir para identificar a gravidade e prever eventos, antes que eles aconteçam, sempre com vistas a segurança dos pacientes hospitalizados, como sugere AL-SHWAHEEN et al. (2022).

A previsão de incidentes clínicos com o paciente, como as deteriorações clínicas e mortalidade, devem ser ampliadas nos próximos anos, como reforçam **BELL et al (2021)**.

Recentemente, **STEINBERG et al. (2021)** demonstraram que muitos estudos estão sendo conduzidos para otimização de predições de eventos com base em Registros Eletrônicos de Saúde (RES), mediados por *Machine Learning* e *Deep Learning* (Aprendizado de Máquina e Aprendizado Profundo).

Hoje, temos disponíveis desde base de dados privadas em instituições e organizações de saúde até em bases públicas para pesquisas, espalhadas por diversos continentes. No entanto, **BERKOWITZ et al. (2024)**, apontam que a descoberta do conhecimento a partir de Registros Eletrônicos de Saúde (RES), não é uma atividade trivial, são intervenções especializadas e não podem ser transferidas e identificadas manualmente, além de questões de ordem ética e de proteção de dados sensíveis, que devem ser consideradas, enfatiza **RUIZ et al. (2022)** e **KORACH et al. (2020)**.

O crescente volume de dados clínicos nos Prontuário Eletrônico do Paciente (PEP), conhecidos como Registros Eletrônicos de Saúde (RES) ou *Electronic Health Records* (EHRs), se caracterizam como registros longitudinais de informações de saúde dos pacientes, como sugerem **MALM-NICOLAISEN et al. (2019)**, ainda que incluídos os dados estruturados como: demográficos, sinais vitais, exames de laboratório e não estruturados como: textos livres, anotações, relatórios técnicos e evoluções do paciente por equipes multidisciplinares, onde sugerem que nos dias atuais, a manipulação destes dados, seja por meio de técnicas avançadas de inteligência artificial (IA), como a limpeza, tratamento e o treinamento com o aprendizado de máquina (**CHIU et al. 2023**).

Atualmente é possível utilizar modelos de protocolos que podem observar, monitorar deteriorações clínicas e riscos de mortalidade em pacientes hospitalizados, segundo **MENG et al. (2022)**. Em sua maioria, observamos modelos manuais, que podem sugerir a tomada de decisão de profissionais de saúde, como enfermeiros, médicos, fisioterapeutas, nutricionistas, psicólogos entre outros. Estes sistemas de identificação de deteriorações clínicas, são hoje, em sua maioria, muito específicos e focados em aspectos direcionados a condições com resultados limitados (**ROMERO-BRUFU et al. 2021**).

**MISWAN et al. (2021)** fazem previsões que estamos em uma tendência global, onde os registros manuais em saúde, sejam brevemente substituídos apenas por Registros Eletrônicos de Saúde (RES). Desta forma, deveremos observar propostas de *insights* preditivos cada vez

mais constantes no ambiente de assistência em saúde. Com o predomínio do uso dos RES, se observa a oferta de oportunidades para a comunidade científica com uso da mineração de dados em saúde, inteligência artificial generativa e aprendizado profundo de dados. **(SUNDERMANN et al., 2019 e VAN DIJK et al., 2021 e CHO et al., 2020)**

No Brasil, em 2024, mais de 25 milhões de pacientes visitaram os serviços públicos de saúde, sejam na atenção primária, onde acontece a promoção a saúde, seja na atenção secundária, onde acontece a média complexidade como os prontos socorros, pronto atendimentos, ambulatórios ou na atenção terciária, onde ocorre os atendimentos nos ambientes hospitalares de alta complexidade. O maior volume destes registros em prontuário, estão em prontuários eletrônicos e se transformam em Registros Eletrônicos de Saúde (RES) e consequente, grande volume de dados, que podem ser utilizados pelos próprios pacientes. Também, para organização da sistematização da prestação do cuidado, para o planejamento epidemiológico das autoridades sanitárias e de saúde, para o ensino e pesquisa e para o âmbito jurídico. **(OLIVEIRA et al., 2025)**

O Prontuário Eletrônico do Paciente (PEP), é um recurso digital, evoluído do tradicional conhecido, prontuário manual do paciente. Este recurso deve conter todo o histórico de saúde de um paciente, incluindo informações como: Dados pessoais, histórico de saúde, dados cirúrgicos e clínicos, resultados de exames, medicações prescritas, alergias, procedimentos realizados, evoluções e anotações dos profissionais de saúde. **(COUTINHO et al., 2025 e ARMAIGNAC et al., 2025)**

Em recente publicação de **CHEN et al. (2024)**, se destaca cada vez mais, o uso do prontuário eletrônico que está sendo incorporado nas práticas organizacionais das instituições e dos profissionais de saúde. Observamos que desde a promoção da saúde, que acontece nas Unidades Básicas de Saúde (UBS), Estratégias de Saúde da Família (ESF) passando pelos Ambulatórios Médicos de Especialidades (AMEs) até no ambiente hospitalar, sejam públicos ou privados, os prontuários e sistemas de registros em saúde para guardar os dados, são organizados em modelos, chamados de arquiteturas estruturadas ou não estruturadas, eletronicamente interoperáveis de maneira a facilitar o acesso à leitura e a transmissão dos dados registrados. **(GROENHOF et al., 2020; DATTA et al., 2022)**

De modo geral, os registros eletrônicos em saúde, são observados em formato de números, imagens ou textos que podem ser explorados por meio de técnicas de inteligência artificial e mineração de dados a depender da necessidade e do tipo de análise necessária. **(HU et al., 2023 e AKEL et al., 2021)**

Os registros eletrônicos em saúde, quando totalmente instalados e em pleno funcionamento, são muito mais do que simples repositórios de dados, uma vez que são registrados em tempo real ou armazenados, centrados sempre no paciente. Neste sentido, **CHOI et al. (2023)** destacam que estes dados podem ter acessos em tempo real ou instantâneos, "quando e onde forem necessários", além de reunir em um só lugar, tudo sobre a saúde do paciente. Afirmações também destacadas por **TANG et al (2020)**.

Com uso do prontuário eletrônico do paciente, podemos automatizar e agilizar tarefas com apoio de sistemas inteligentes, fazendo com que os registros eletrônicos, possam gerar *insights* e previsões de eventos futuros, além de apoiar a organização epidemiológica e tomada de decisões de governança clínica e planejamento estratégico do setor público, da saúde suplementar e do setor privado. (**FANELLI et al. 2023**)

**HERRERO-ZAZO et al (2023)** sugerem que a inteligência artificial (IA) é a capacidade associada a sistemas computacionais, que são capazes de simular a inteligência humana por meio de aprendizado de máquina. A inteligência artificial, com seus algoritmos, é capaz de apoiar médicos, enfermeiros, entre outros profissionais a diagnosticar e planejar cuidados aos pacientes, sejam em situações clínicas ou cirúrgicas com o mapeamento de tendências e padrões que são mais difíceis de serem observados de forma manual. (**GONDIM et al., 2022**).

Com a transformação da saúde digital, podemos utilizar de recursos integrados com dispositivos eletrônicos e os chamados sistemas inteligentes para observar o comportamento dos parâmetros vitais e promover intervenções em tempo hábil, com objetivo de inibir ou ajustar o curso da evolução do quadro clínico dos pacientes (**SHAMOUT et al., 2023**).

Podemos observar nos ambientes de práticas de assistência à saúde, que as rotinas manuais, estão sendo substituídas progressivamente por sistemas de apoio a decisão, com o objetivo de se evitar a dependência direta do controle exclusivo do fator humano dos profissionais de saúde. Os sistemas inteligentes têm se tornado ferramentas cada vez mais importantes na detecção de agravamentos dos pacientes e suas deteriorações. (**HARKANEN et al. 2019 e FANELLI et al., 2023**)

Vale destacar, que por meio de inteligência artificial, conhecer o comportamento progressivo dos parâmetros clínicos dos pacientes, nos possibilita ajustar algoritmos, para que nos mostrem em tempo seguro, o momento crucial de promover intervenções no ajuste fisiológico do corpo humano, a exemplo da abordagem apresentada no estudo de **NGUYEN et al. (2022)**.

Há outros importantes estudos, tais como os trabalhos de **HAMLIN et al. (2023)** e

**PADILHA et al. (2018)**, os quais corroboram o fato de que outros eventos clínicos do tipo metabólicos e neurológicos também são possíveis de deteriorações clínicas e/ou laboratoriais, e podem ser observados de 24 a 48 horas nos sinais e sintomas aparentes.

Diante deste cenário apresentado, e apesar do amplo uso de escalas manuais e outros recursos no ambiente médico-hospitalar, a presente pesquisa considerou a possibilidade de desenvolvimento de um índice clínico-laboratorial, ou seja, um novo indicador composto pelos sinais vitais, por resultados de exames laboratoriais do paciente e por observações clínicas, obtidos a partir de prontuários eletrônicos, a fim de melhor caracterizar a sua possível deterioração, quando internado em uma unidade de terapia intensiva. Foi previsto ainda o uso de tal indicador, calculado ao longo de um dado período, para cada paciente, em conjunto com um algoritmo de aprendizagem de máquina profunda, no sentido de prever a sua evolução temporal (deterioração) e apoiar a decisão dos profissionais de saúde no aludido ambiente.

### 1.1 MOTIVAÇÕES DA PESQUISA

Entre várias motivações da pesquisa, uma das que merece maior destaque, foi a oportunidade de o pesquisador ter vivenciado *in loco*, no ambiente de práticas assistenciais hospitalares, durante mais de 15 anos e observado inúmeras oportunidades de criação de um sistema de apoio a decisão clínica.

As manifestações de profissionais, como enfermeiros e médicos, onde sugerem que deva existir algum recurso, se possível mediado por tecnologias emergentes, para prever antecipadamente eventos de deteriorações clínicas e mortalidade futuras de pacientes hospitalizados.

Também merece valorização para a pesquisa, emerge da vivência prática e científica no cuidado a pacientes adultos hospitalizados, especialmente da observação de que eventos graves de deterioração clínica, muitas vezes, **não ocorrem de forma súbita**, mas são precedidos por alterações progressivas em sinais vitais e exames laboratoriais que passam despercebidas ou são reconhecidas tardiamente na rotina assistencial.

No contexto hospitalar, em especial em enfermarias e unidades de cuidados intermediários, a vigilância clínica depende fortemente da avaliação humana, da experiência profissional e de instrumentos tradicionais de escore clínico. Embora úteis, tais instrumentos apresentam limitações importantes, como rigidez de regras, sensibilidade reduzida a mudanças sutis e dificuldade de adaptação a diferentes perfis de pacientes. Essa realidade motivou a busca

por abordagens mais flexíveis, dinâmicas e orientadas por dados, capazes de refletir a complexidade do estado clínico ao longo do tempo.

É crescente digitalização da assistência à saúde, uma ampla adoção de Prontuários Eletrônicos de Saúde que resultou na geração contínua de grandes volumes de dados clínicos longitudinais. Entretanto, esses dados permanecem, em grande parte, subexplorados do ponto de vista preditivo, sobretudo em aplicações que integrem informações clínicas e laboratoriais de forma temporal e contínua. Essa constatação reforçou o interesse em investigar como técnicas avançadas de ciência de dados e aprendizagem profunda podem transformar registros rotineiros em ferramentas efetivas de apoio à decisão clínica.

Do ponto de vista científico, a pesquisa é motivada pela necessidade de avançar além de modelos isolados, explorando **arquiteturas híbridas de aprendizado profundo** que combinem diferentes capacidades analíticas. A integração de redes convolucionais e recorrentes permite capturar tanto relações complexas entre variáveis clínicas quanto dependências temporais de curto e longo prazo, características fundamentais das séries temporais em saúde. Essa abordagem representa uma oportunidade de inovação metodológica na predição de deterioração clínica.

Além disso, a pesquisa é motivada por um compromisso com a qualificação do cuidado em saúde, especialmente no contexto do sistema público. Ferramentas preditivas bem validadas podem apoiar equipes multiprofissionais, favorecer intervenções precoces, reduzir eventos adversos evitáveis e contribuir para uma assistência mais segura, eficiente e equitativa. Para a enfermagem, em particular, tais soluções representam um avanço no suporte à tomada de decisão clínica baseada em dados.

Por fim, esta pesquisa é motivada pelo desejo de contribuir para o avanço científico nacional em informática em saúde, fortalecendo a produção de conhecimento, a inovação tecnológica e o desenvolvimento de soluções alinhadas às realidades assistenciais e epidemiológicas locais. Ao integrar prática clínica, ciência de dados e aprendizagem profunda, o estudo busca consolidar uma abordagem interdisciplinar capaz de gerar impacto científico, assistencial e social.

## 1.2 JUSTIFICATIVAS DA PESQUISA

O estudo de **DONG ZHANG et al. (2020)**, projetou uma combinação de dados estruturados e não estruturados para modelos também preditivos com abordagem de aprendizagem profundo, conseguiu-se demonstrar riscos de mortalidade, readmissão hospitalar até 30 dias e previsões de longa duração, onde também foi usada a base (MIMIC-III), *US Medical Information mart for Intensive Care*, de pacientes adultos com uma amostra de 40.000 pacientes. Nesta pesquisa, foram utilizadas duas redes neurais profundas multimodais como *Fusion - LSTMs* e *Fusion - CNNs* nos dados estruturados e nos dados não estruturados as técnicas de Processamento de Linguagem Natural.

Já as pesquisas de **SHRESTHA et al. (2021)** e **CHIU et al., (2023)**, utilizaram em seus desenhos de pesquisa, técnicas de Processamento de Linguagem Natural (PNL), Técnicas de Reconhecimento de Entidades Nomeadas (NER), Redes Neurais, apoiadas por Modelos Pré-Treinados amplamente reconhecidos na área de saúde para identificação de previsões de mortalidade.

Outro estudo que merece destaque, são os resultados descritos na publicação de **BAKER et al. (2020)**, que explora uma proposta para a previsão contínua e automática do risco de mortalidade em Unidades de Terapia Intensiva (UTIs). Os autores desenvolveram um modelo de rede neural híbrida que combina camadas convolucionais (*CNN*) com memória bidirecional de longo prazo (*BiLSTM*) para prever a mortalidade com base em sinais vitais ao longo de um período de 24 horas. O modelo foi avaliado usando dados do banco de dados MIMIC-III e demonstrou forte desempenho em comparação com esquemas de última geração para previsão de mortalidade.

Entende-se assim a necessidade de explorar e propor novas alternativas, com ampliação dos dados de entrada, ou seja, os *inputs*, considerando que os registros eletrônicos de saúde, sejam demográficos, sinais vitais, exames laboratoriais e textos livres como propõe **AUNG et al. (2021)**. Desta forma, observa-se um leque de oportunidades para estratificar um robusto volume de registros eletrônicos, sendo produzidos continuamente.

No que se refere aos incidentes e a mortalidade hospitalar, nos deparamos com uma oportunidade tecnológica, ainda pouco explorada por profissionais de saúde, onde encontra-se maneiras de prever acontecimentos não desejados em pacientes hospitalizados, com isso, pode-se também reduzir o desconforto de familiares e ao sistema de saúde, além de ser uma contribuição científica para a comunidade acadêmica.

No ambiente hospitalar, os cuidados aos pacientes não são atividades simples, com diversos aspectos que podem interferir no prognóstico de um paciente instável. Desta forma, conseguir prever e antever acontecimentos clínicos, pode ser alternativas cruciais para definir o planejamento de intervenções e condutas, a fim de evitar as intercorrências, que podem colocar o paciente, os profissionais e as instituições sobre risco.

A previsão de risco de mortalidade em pacientes adultos hospitalizados, constitui uma intervenção essencial para o suporte à decisão clínica, contribuindo para a implementação de intervenções precoces e otimização de recursos hospitalares.

Com os avanços da inteligência artificial (IA), sobretudo na área de IA generativa, busca-se desenvolver a construção de modelos preditivos de aprendizagem profundo mais precisos e interpretáveis, capazes de integrar diferentes fontes de dados. A aplicação de IA generativa neste contexto, permite a elaboração de índices de risco dinâmicos, contextualizados e adaptativos, que podem atuar como suporte de decisão para enfermeiros e médicos.

Este estudo, justifica-se pela necessidade de desenvolver um sistema preditivo fundamentado em aprendizagem de máquina profundo, capaz de processar dados clínicos heterogêneos de registros eletrônicos em saúde, contribuindo para a redução da mortalidade hospitalar e para a elevação da qualidade do cuidado. Além disso, busca promover inovação na integração de tecnologias de ponta na rotina hospitalar, fomentando uma abordagem baseada em evidências para o gerenciamento de riscos clínicos.

**TABELA 1 – PUBLICAÇÕES QUE NORTEARAM A PESQUISA**

<b>TÍTULO</b>	<b>AUTOR/ANO</b>	<b>BASES</b>	<b>TEMA/DESFECHO</b>	<b>ARQUITETURA</b>	<b>CITAÇÕES</b>
MIMIC-III, a freely accessible critical care database	Johnson et al., 2016	MIMIC-III	BASE DE DADOS	BASE DE DADOS	9969
Multitask learning and benchmarking with clinical time series data	Harutyunyan et al., 2019	MIMIC-III	Deteriorações Clínicas	CNN/LSTM)	1320
Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory	Lin et al., 2019	MIMIC-III	Deteriorações Clínicas	RNN/LSTM	251
Early prediction of sepsis in the ICU using machine learning	Moore et al., 2021	MIMIC	Lesão renal aguda	CNN/LSTM)	229
Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques	Miao et al., 2020	MIMIC	Tempo de permanência	CNN/LSTM	153
End-to-end blood pressure prediction via fully convolutional networks	Baek, Jang, & Yoon, 2019	MIMIC	Modelagem fisiológica contínua	CNN/LSTM)	88
Continuous and automatic mortality risk prediction using vital signs in the intensive care unit: A hybrid neural network approach	Baker et al., 2020	MIMIC-III	Mortalidade em UTI	CNN-LSTM	58
Foreinsight – A generative pretrained transformer for modelling of patient timelines using electronic health records: a retrospective modelling study	Kraljevic Z et al, 2024	MIMIC-III	Deteriorações Clínicas	TRANSFORMERS	54
PulseDB: A large, cleaned dataset based on MIMIC-III and VitalDB for benchmarking cuff-less blood pressure estimation methods	Wang et al., 2022	MIMIC-III	Múltiplos desfechos clínicos	CNN-LSTM	20
A novel deep learning algoritmo for real-time prediction of clinical deterioration in the emergency department for a multimodal clinical decision support system	Choi A et al, 2024	USME	Deteriorações Clínicas	TRANSFORMERS	19
A novel and reliable framework of patient deterioration prediction in intensive care unit based on long short-term memory-recurrent neural network	Tarik I et al, 2021	MIMIC-III	Deteriorações Clínicas	RNN/LSTM	18
A computationally efficient CNN-LSTM neural network for estimation of blood pressure from features of electrocardiogram and photoplethysmogram waveforms	Baker et al., 2022	MIMIC-III	Mortalidade	CNN-LSTM	17
Multimodal temporal-clinical note network for mortality prediction	Yang et al., 2021	MIMIC-III	Readmissão em UTI	CNN (texto) + LSTM	16
Combined deep CNN–LSTM network-based multitasking learning architecture for noninvasive continuous blood pressure estimation using difference in ECG-PPG features	Jeong et al., 2021	MIMIC-III	Descompensação / Mortalidade	CNN-LSTM	15
Classification of blood pressure levels based on photoplethysmogram and electrocardiogram signals with a concatenated convolutional neural network	Fuadah & Lim, 2022	MIMIC	Deteriorações Clínicas	CNN-LSTM	11
Acute complication prediction and diagnosis model based on MIMIC-IV (CLSTM-BPR)	Chen et al., 2024	MIMIC-IV	Deteriorações Clínicas	CLSTM (CNN + LSTM)	7
Predicting decompensation risk in intensive care unit patients using machine learning	Aikodon et al., 2024	MIMIC-IV	Readmissão em UTI	CNN-LSTM	4
Deep learning model for predicting sepsis using MIMIC-III	Rafei & Mello, 2021	MIMIC-III	Sepse precoce	CNN/LSTM	0

O estudo de **BAKER et al. (2020)** apresenta uma abordagem híbrida de rede neural para a predição contínua e automática de risco de mortalidade em pacientes na Unidade de Terapia Intensiva (UTI), utilizando sinais vitais. Os achados destacam a eficácia dos modelos computacionais em prever eventos adversos, o que substancia a relevância do uso de algoritmos de aprendizado de máquina nesta atual proposta de pesquisa. Essa abordagem orienta a construção de um novo índice de predição de deteriorações clínicas, fornecendo uma estrutura comprovada que pode ser adaptada e aprimorada para incluir variáveis adicionais pertinentes ao meu estudo.

No trabalho de **RAFIEI E MELLO (2021)**, um modelo de aprendizado profundo foi desenvolvido para prever sepse utilizando dados do MIMIC-III. A pesquisa demonstra como as técnicas de aprendizado profundo podem ser aplicadas a grandes conjuntos de dados clínicos, ressaltando a importância de dados robustos e de alta qualidade. Essa fundamentação é crucial para esta atual proposta de pesquisa, pois planeja-se utilizar dados semelhantes para treinar um modelo que antecipe deteriorações clínicas, refletindo a importância da precisão na predição do estado do paciente.

O artigo de **LIN et al. (2019)** investiga a readmissão não planejada na UTI por meio de redes neurais recorrentes com memória de longo prazo (LSTM). Os resultados indicam que modelos que consideram a temporalidade dos dados podem melhorar a precisão das predições. Essa perspectiva é particularmente relevante para um novo índice preditivo, já que a deterioração clínica laboratorial muitas vezes mostra padrões temporais que precisam ser considerados para intervenções antecipadas.

**AIKODON et al. (2024)** exploram a predição do risco de descompensação em pacientes de UTI usando machine learning. O estudo estabelece a importância da modelagem preditiva no ambiente crítico de cuidados. As metodologias e conceitos discutidos neste artigo são diretamente aplicáveis ao desenvolver um novo índice de predição, pois enfatizam a necessidade de intervenções precoces e a utilidade de algoritmos para auxiliar nessa tarefa.

**YANG et al. (2021)** desenvolvem uma rede multimodal para a predição de mortalidade, incorporando notas clínicas temporais. A integração de dados clínicos não estruturados na predição é uma abordagem poderosa que pode enriquecer a capacidade de meu modelo. Ao empregar técnicas semelhantes, pode-se melhorar a sensibilidade e especificidade de um novo índice de deterioração clínica laboratorial.

O estudo de **CHEN et al. (2024)** apresenta um modelo de predição de complicações agudas com base em dados do MIMIC-IV. A pesquisa demonstra a evolução das abordagens de

aprendizado de máquina em ambientes hospitalares, fornecendo um contexto contemporâneo para este novo trabalho. A utilização de uma base de dados semelhante reforça a validação de um novo índice, alinhando-se com as melhores práticas na área.

No artigo de **JEONG et al. (2021)**, é explorada uma arquitetura de aprendizado profundo que combina redes neurais convolucionais (CNN) e LSTM para estimar continuamente a pressão arterial a partir de características de eletrocardiogramas. A relevância deste estudo está na abordagem de multitarefa que pode ser adaptada para esta atual proposta de pesquisa, utilizando dados vitais para prever deteriorações clínicas. A metodologia proposta oferece uma base sólida para a integração de diferentes tipos de dados, permitindo a utilização de múltiplas entradas para aprimorar a precisão de um novo índice preditivo.

**BAKER et al. (2022)** apresentaram uma rede CNN-LSTM computacionalmente eficiente para a estimativa de pressão arterial utilizando dados de eletrocardiogramas. A eficiência computacional é um fator importante na implementação clínica de modelos preditivos. Este artigo é relevante para esta atual proposta de pesquisa, pois fornece *insights* sobre como otimizar modelos para uso em tempo real, o que é crucial para a aplicação prática do meu índice preditivo em ambientes clínicos dinâmicos.

O estudo de **FUADAH E LIM (2022)** aborda a classificação dos níveis de pressão arterial com base em sinais vitais, utilizando redes neurais convolucionais concatenadas. A pesquisa demonstra como diferentes sinais fisiológicos podem ser combinados para melhorar a precisão da predição clínica. Isso reflete o potencial de integrar múltiplas variáveis em um novo índice de deterioração clínica laboratorial, aumentando a robustez e a confiabilidade das predições.

**WANG et al. (2023)** desenvolveram o PulseDB, um grande conjunto de dados limpos baseado no MIMIC-III e VitalDB para a avaliação de métodos de estimativa de pressão arterial sem manguito. A disponibilidade de conjuntos de dados de alta qualidade é essencial para o treinamento eficaz de modelos preditivos. Este artigo é fundamental para esta atual proposta de pesquisa, pois pretende-se utilizar conjuntos de dados semelhantes para garantir que o índice de deterioração clínica laboratorial se baseie em dados confiáveis e representativos.

A publicação de **JOHNSON et al. (2016)** apresenta o MIMIC-III, um banco de dados de cuidados críticos amplamente acessível, que tem sido uma fonte valiosa para muitos estudos em aprendizado de máquina na saúde. A utilização desse banco de dados nesta pesquisa poderá viabilizar acesso a um vasto conjunto de dados clínicas, permitindo a validação e teste do novo índice preditivo com base em um padrão estabelecido de pesquisa.

O trabalho de **HARUTYUNYAN et al. (2019)** foca no aprendizado multitarefa com dados clínicos em séries temporais. Essa abordagem é pertinente à esta pesquisa, pois sugere que a modelagem simultânea de diferentes resultados clínicos pode levar a melhores previsões. Integrar essa estratégia ao novo índice de deterioração, poderá melhorar a identificação precoce de múltiplas complicações em pacientes hospitalizados.

A pesquisa de **MOORE et al. (2021)** avalia a previsão precoce de sepse no ambiente da UTI utilizando técnicas de aprendizado de máquina, incluindo CNN-LSTM. A análise demonstrou que intervenções precoces podem reduzir a mortalidade. Este estudo reforça a importância de prever deteriorações clínicas, apoiando a adoção do novo índice de previsão para facilitar ações tempestivas em pacientes em risco de complicações.

**MIAO et al. (2020)** apresentam um método para a medição contínua da pressão arterial a partir de um sinal de ECG unidimensional, utilizando técnicas de aprendizado profundo. A inovação deste trabalho destaca como técnicas avançadas podem aprimorar a monitorização dos sinais vitais. Essa perspectiva é valiosa para esta atual proposta de pesquisa, pois a precisão na avaliação dos sinais vitais é fundamental para o sucesso do novo índice preditivo de deterioração clínica laboratorial.

**BAEK et al. (2019)** discutem a previsão da pressão arterial usando redes neurais convolucionais totalmente convolucionais. Este estudo oferece modelos end-to-end que podem ser implementados em *pipelines* clínicos. A adoção de suas técnicas pode facilitar a execução do índice preditivo em tempo real, permitindo que os profissionais de saúde respondam rapidamente às alterações críticas nas condições dos pacientes. A contribuição deste artigo para a eficiência dos algoritmos de previsão é diretamente aplicável à construção do novo modelo, que almeja integrar dados vitais de maneira fluida e em tempo real.

O estudo de **CHOI et al. (2024)** apresenta um algoritmo inovador para a previsão em tempo real da deterioração clínica no departamento de emergência, utilizando um sistema de suporte à decisão clínica multimodal. Este trabalho é altamente relevante para esta atual proposta de pesquisa, pois destaca a importância de sistemas que podem integrar diferentes modalidades de dados clínicos para melhorar a previsão e a resposta em situações críticas. A abordagem proposta por **CHOI et al. (2024)** irá inspirar a estrutura e a funcionalidade do novo índice preditivo, permitindo uma melhor resposta em cenários clínicos dinâmicos.

A análise das publicações mencionadas, fornece uma base para nortear para esta atual proposta de pesquisa ao demonstrar as diversas abordagens e metodologias aplicáveis no campo da predição de deteriorações clínicas. Cada artigo contribui de maneira única para a construção de um novo índice preditivo, incorporando técnicas de aprendizado de máquina, dados clínicos relevantes e abordagens inovadoras. Essa fundamentação não apenas valida a importância do deste trabalho, mas também oferece diretrizes e inspiradoras para seu desenvolvimento, visando à melhoria da assistência clínica e à saúde dos pacientes hospitalizados.

### 1.3 IDENTIFICAÇÃO DE LACUNAS DA PESQUISA

Observou-se neste estudo, que as lacunas da pesquisa, envolvem a deterioração clínica do paciente, a integração de dados não estruturados e estruturados; aplicação de inteligência artificial generativa na saúde; validação do modelo com dados reais; suporte a decisão clínica personalizada, impactos na mortalidade e na gestão de recursos.

- **Deterioração Clínica do paciente:** Apesar dos avanços no monitoramento clínico e na digitalização dos prontuários eletrônicos de saúde, a deterioração clínica em pacientes hospitalizados ainda representa um desafio significativo para os sistemas de saúde. A literatura demonstra progressos no desenvolvimento de escores de alerta precoce e, mais recentemente, de modelos baseados em aprendizado de máquina; entretanto, persistem importantes lacunas. Muitos estudos concentram-se em populações específicas ou em cenários altamente controlados, com limitada validação externa e reduzida aplicabilidade em contextos reais de prática clínica. Além disso, ainda é insuficiente a integração de dados clínicos e laboratoriais longitudinais de forma dinâmica, capaz de capturar a evolução temporal do paciente. Soma-se a isso a carência de modelos interpretáveis, interoperáveis com sistemas hospitalares e validados em diferentes realidades epidemiológicas, especialmente em países de média renda. Dessa forma, permanece uma lacuna relevante no desenvolvimento e validação de abordagens preditivas robustas, baseadas em dados do mundo real, que apoiem a identificação precoce da deterioração clínica e subsidiem intervenções oportunas na prática assistencial. (PADILHA R. et al, 2018 e HOLSWORTH L. et al, 2021).
- **Integração de dados não estruturados ou estruturados:** Ainda há um desafio significativo na integração eficiente de diferentes tipos de dados clínicos, como textos

livres, sinais vitais e exames laboratoriais, para gerar modelos preditivos robustos e interpretáveis. A integração de distintas fontes de dados é essencial para melhorar a precisão de modelos preditivos em saúde. Dados estruturados, como exames laboratoriais e sinais vitais, são facilmente manipulados e analisados, enquanto dados não estruturados, como textos clínicos livres, representam uma parcela significativa de informações clínicas complementares e ricas. Ainda assim, a integração eficiente desses tipos de dados apresenta desafios técnicos e metodológicos, incluindo a heterogeneidade de formatos, a necessidade de técnicas avançadas de processamento de linguagem natural (PLN) e de interoperabilidade entre sistemas. (MIOTTO et al. 2016).

As abordagens atuais, como integração via representações vetoriais e multimodal *deep learning*, mostram potencial, porém ainda enfrentam limitações em escalabilidade e interpretabilidade (MIOTTO et al., 2018). Portanto, há um avanço necessário na pesquisa de métodos que possam unir efetivamente esses dados para gerar modelos preditivos mais robustos e interpretáveis. Em resumo, deve-se desenvolver e validar métodos de integração que sejam escaláveis, capazes de lidar com o volume e a heterogeneidade dos dados clínicos.

- **Aplicação de IA preditiva na saúde:** Embora a IA preditiva tenha sido amplamente explorada em outras áreas, sua utilização específica para previsão de desfechos clínicos em ambientes hospitalares, especialmente na incorporação de dados não estruturados, permanece limitada e pouco evidenciada. A IA generativa, que inclui técnicas como modelos de linguagem natural avançados (ex.: GPT), redes neurais convolucionais (CNN), entre outras, tem mostrado grande potencial em tarefas de geração de texto, síntese de informações e suporte à tomada de decisão em diversas áreas, como linguagem e criatividade (SAI S. et al., 2024). No entanto, sua aplicação em saúde clínica ainda está em fase inicial, especialmente na previsão de desfechos clínicos com incorporação de dados não estruturados (ZHANG D et al., 2020).

Estudos recentes como o de PETERS et al (2020) sugerem que a IA preditiva pode auxiliar na elaboração de relatórios, na extração automática de informações de textos clínicos e na previsão de eventos adversos a partir de dados heterogêneos, mas sua utilização para prever desfechos complexos ainda carece de validação empírica robusta e de avaliação de confiabilidade na prática clínica.

Portanto, em síntese, deve-se investigar e validar o uso de IA generativa especificamente para previsão de desfechos clínicos, com foco na incorporação de dados não estruturados em ambientes hospitalares.

- **Validação de Modelos com dados reais:** Grande parte dos modelos preditivos desenvolvidos ainda não foram validados de forma ampla em contextos clínicos reais, limitando sua aplicabilidade prática e sua aceitação pelos profissionais de saúde. Muitos modelos preditivos desenvolvidos em estudos de academia permanecem em fase de validação interna ou em ambientes controlados, e raramente são testados em clínicas ou hospitais em prática, limitando sua adoção (STEYERBERG et al., 2019). A validação externa e a implementação em ambientes clínicos reais são cruciais para verificar o modelo e sua efetividade na rotina assistencial. Estudos como de RAJKOMAR et al. (2019), mostram que a introdução de modelos em contextos reais, frequentemente apresenta desafios adicionais, como a variabilidade dos dados, resistência à mudança por parte dos profissionais e dificuldades na integração com os fluxos de trabalho clínico. Assim, há necessidade de pesquisas que foquem não apenas no desenvolvimento ou na aplicação, mas também na validação prática e na aceitação por profissionais de saúde. Desta forma, este estudo deverá produzir experimentos e testes com base em dados reais, extraídos de uma base de dados de pacientes hospitalizados, sendo uma parte aproximada de 80% para treinamento dos modelos e outra parte de 20% para testes e validações das métricas. Sabe-se que, estudos de validação externa de modelos preditivos em diferentes ambientes clínicos, que avalie fatores de aceitação e impacto real na prática, são recomendados.
- **Suporte à decisão clínica personalizada:** Há uma lacuna na pesquisa sobre como esses modelos podem oferecer recomendações interpretáveis e personalizadas, para facilitar a tomada de decisão por enfermeiros e médicos de forma efetiva e segura. A personalização do suporte à decisão clínica, que leva em consideração as múltiplas variáveis específicas de cada paciente, é uma meta central na ciência médica de precisão. Modelos interpretáveis e recomendatórios que possam apresentar não somente uma previsão, mas também uma explicação clara, são essenciais para promover a confiança e facilitar a adesão por parte de médicos e de enfermeiros (DOSHI-VELEZ & KIM, 2017). Apesar dos avanços em aprendizado de máquina, há uma lacuna na literatura sobre como transformar modelos preditivos complexos em ferramentas de

suporte à decisão verdadeiramente utilizáveis, que forneçam recomendações clínicas interpretáveis, seguras e personalizadas (CARUANA et al., 2015).

Neste sentido, importante destacar que em contextos de alta complexidade as muitas variáveis clínicas são fatores a se considerar na análise do cenário de cada paciente. Com isso, faz-se desenvolver e testar algoritmos de suporte à decisão médica, que sejam interpretáveis, clinicamente relevantes e capazes de oferecer melhores previsões possíveis.

- **Impacto na mortalidade e na gestão de recursos:** Ainda há poucas evidências com métricas de desempenho de implementação com ferramentas preditivas, que demonstrem redução efetiva da mortalidade hospitalar e na otimização do uso de recursos de cuidado. (HASHIR et al. 2020). Ainda assim, evidências empíricas de que esses sistemas promovem melhorias nos resultados, permanecem escassas ou inconclusivas (CHEN J et al., 2023). A complexidade da implementação, resistência dos profissionais e fatores organizacionais, dificultam a avaliação de impacto real. Deste modo, vale refletir, com quais estratégias de implementação e avaliação, podem evidenciar de forma convincente o impacto de sistemas preditivos na mortalidade e na gestão de recursos hospitalares.

### **Relevância e Ineditismo**

A relevância deste estudo reside na proposição de um índice clínico-laboratorial inédito, que preenche uma lacuna crítica nos sistemas de pontuação tradicionais ao integrar variáveis biológicas dinâmicas a sinais vitais convencionais. O ineditismo da abordagem sustenta-se na utilização de um modelo híbrido de aprendizagem profunda como CNN - LSTM, capaz de extrair padrões complexos e dependências temporais em grandes volumes de dados de pacientes hospitalizados. Diferente de modelos unidimensionais ou índices puramente clínicos, esta pesquisa inova ao oferecer uma ferramenta de previsão de alta fidelidade que potencializa a antecipação da deterioração assistencial, permitindo uma intervenção precoce e personalizada que visa, fundamentalmente, a redução da mortalidade e a otimização de recursos críticos no ambiente hospitalar. Além de propor um modelo híbrido, treinado, validado e testado em duas importantes bases de dados de pacientes reais como MIMIC III e MIMIC IV.

## 1.4 QUESTÃO DE PESQUISA

"Como um índice numérico formado por informações clínicas e laboratoriais, a partir de dados de registros eletrônicos de saúde, e o seu uso com um algoritmo de aprendizagem de máquina, pode ser capaz de oferecer suporte às decisões clínicas tomadas por enfermeiros e médicos, para avaliar a deterioração em pacientes adultos hospitalizados?"

## 1.5 OBJETIVOS

**1.5.1 Objetivo** do estudo é desenvolver e testar um índice composto por dados clínicos e de laboratório para a predição de deterioração em pacientes adultos hospitalizados com o uso de algoritmos de aprendizagem de máquina.

**1.5.2 Os objetivos específicos** do estudo, são:

- Criar um índice para a predição futura da deterioração clínica e laboratorial de pacientes adultos, a partir de sinais vitais e exames laboratoriais de cada paciente, considerando as variáveis estatisticamente mais relevantes;
- Criar um banco de dados que represente a variação temporal do índice para cada paciente;
- Desenvolver, treinar e validar ao menos um modelo de aprendizado de máquina, a partir de um primeiro banco de dados criado, usando o dataset público conhecido como MIMIC-III, capaz de prever a deterioração clínica de um paciente, a partir do uso do índice em intervalos de tempos passados de deteriorações clínicas;
- Testar o desempenho do modelo de aprendizagem de máquina treinado, a partir de casos de pacientes não empregados na fase de treino e validação, bem como a partir de um segundo banco de dados usando o dataset público conhecido como MIMIC-IV.

## 1.6 DELIMITAÇÕES DO TEMA DA PESQUISA

A proposta de estudo da pesquisa, deverá focar em desenvolver e propor um novo índice numérico, entre 0 - 100, como um score que envolve análise multivariada de informações clínicas, onde será demonstrada as nuances de entrada de informações dos mesmos pacientes mapeados do *datasets* encontrados e disponíveis dentro da base MIMIC-III (*US Medical Information mart for Intensive Care*). Todas as entradas de dados, observados nos registros eletrônicos de saúde, serão extraídos de pacientes reais, onde o índice deverá resultar em um único valor saída de 0 - 100, compilado com seus respectivos pesos de relevância a depender da condição de cada variável, estabelecida pelo algoritmo de treinamento utilizado ao longo da pesquisa: Primeira variável: “Dados demográficos do paciente”, Segunda variável: “Sinais vitais do paciente”, Terceira variável: “Resultados laboratoriais de análises clínicas”.

Vale destacar que, embora o prontuário eletrônico do paciente, com seus registros, dispõe tradicionalmente de um volume muito grande de informações coletadas. Optou-se por não utilizar resultados de laudos e imagens de exames realizados, informações sobre medicamentos, alguns resultados de escalas e índices aplicados no paciente, como *Braden*, *Morse*, *Mews*, *News*, *Pews*, *Meows*, *Sofa*, *Qsofa*, entre outros, em razão da indisponibilidade regular na base de dados pesquisada.

Outra questão a ser destacada, é que as bases de dados MIMIC (III e IV), utilizadas nesta pesquisa, foram escolhidas por razões éticas, uma vez que os dados são todos anonimizados não sendo possível, portanto, identificar o indivíduo da população de pesquisa. Uma questão, que de certa forma, protege indivíduos e seus dados sensíveis expostos. Garantindo preceitos éticos aos seres humanos.

O presente desenho de pesquisa, também optou em delinear o estudo apenas em dados de pacientes adultos que tiveram desfechos de altas ou óbitos durante a internação.

## 1.7 ORGANIZAÇÃO DO TRABALHO DE PESQUISA

Esta pesquisa está organizada em 7 capítulos, sendo na Introdução, onde são explorados a relevância do contexto da pesquisa, orientações sobre prontuário eletrônico do paciente, registros eletrônicos de saúde, deteriorações clínicas em pacientes, justificativas da pesquisa, motivações, limitações do estudo e objetivos gerais e específicos; revisão da literatura e fundamentação teórica; materiais e métodos; resultado preliminares; Conclusões; Anexos e referências utilizadas na construção da pesquisa.

## 2.0 REVISÃO DA LITERATURA E FUNDAMENTAÇÃO TEÓRICA

Este capítulo procurou fundamentar os conceitos e tecnologias que sustentam o desenvolvimento do modelo preditivo baseado em inteligência artificial (IA) para previsão de riscos de deteriorações clínicas e mortalidade em pacientes adultos hospitalizados, com foco na integração de dados clínicos heterogêneos, sinais vitais e exames laboratoriais. Serão abordados aspectos dos registros eletrônicos de saúde, modelos preditivos em saúde, e as Redes Neurais LSTM (*Long Short-Term Memory*) que é uma arquitetura de rede neural recorrente (RNN). A rede CNN (*Convolutional Neural Network*) ou rede neural convolucional é um tipo de rede neural artificial profunda, especializada em analisar dados com estrutura de grade, como imagens, aprendendo padrões de forma hierárquica, semelhante ao córtex visual humano, sendo algoritmos de *DEEP LEARNING* (Aprendizagem de máquina profundo) utilizado principalmente com o desenvolvimento técnico do algoritmo para aplicação clínica.

A *deterioração clínica* por definição, segundo **DARYL JONES et al, (2013)**, pode ser entendida como a “transição de um paciente de um estado fisiológico relativamente estável para um estado de piora aguda, caracterizada por alterações em sinais vitais e parâmetros fisiológicos que aumentam significativamente o risco de complicações graves e mortalidade. Estudos clássicos destacam que esta deterioração frequentemente precede eventos adversos críticos e pode estar associada a múltiplos fatores intrínsecos e extrínsecos ao paciente” .

A *predição em saúde* é definida por **MALYCHA J. et al, (2022)** com o uso de “dados clínicos e métodos quantitativos para estimar a probabilidade de eventos futuros de saúde, como a deterioração clínica em pacientes hospitalizados. As abordagens incluem desde escores clínicos validados, que somam sinais vitais e parâmetros laboratoriais para estimar risco, até métodos de aprendizado de máquina e inteligência artificial que aprendem padrões complexos nos dados para prever desfechos clínicos”.

As *séries temporais em saúde*, segundo **ZAMMEL Z et al, (2024)** são “coleções de observações quantitativas registradas sequencialmente ao longo do tempo, nas quais cada dado reflete um evento, medida ou indicador de interesse em momentos distintos, mantendo a ordem cronológica como elemento central de análise.”.

## 2.1 A DETERIORAÇÃO CLÍNICA DE PACIENTES EM UTI E O USO DE INDICADORES PARA A SUA AVALIAÇÃO

A deterioração clínica é um fenômeno dinâmico caracterizado pelo declínio agudo ou subagudo do estado fisiológico de um paciente, frequentemente culminando em eventos adversos graves, até com parada cardiorrespiratória (PCR), admissão não planejada em Unidade de Terapia Intensiva (UTI) ou desfecho de óbito. (AKEL et al, 2025 e ALSHWAHHEEN et, al, 2022). Diferente de um evento súbito e imprevisível, a deterioração clínica é, em sua condição, um processo evolutivo que oferece janelas de oportunidades para intervenção, uma vez que alterações nos sinais vitais e exames de laboratório, precedem o colapso catastrófico em até 24 horas na maioria dos casos hospitalares. (BLYTHE et al, 2022 e BOULLTSAKIS et al, 2022).

O centro desencadeador da deterioração clínica, reside na falha dos mecanismos compensatórios homeostáticos. Quando um evento clínico (seja ele infeccioso, isquêmico ou traumático) sobrecarrega a reserva funcional, o organismo manifesta alterações precoces, principalmente no sistema neurológico, respiratório e cardiovascular (BREKKE et al, 2019, CHOI et al, 2023, FU LH, et al, 2021).

### Sistemas de Alertas Precoces (Early Warning Scores)

Historicamente, a resposta à deterioração clínica era reativa. A introdução dos *Early Warning Scores* (EWS), como o *Modified Early Warning Score* (MEWS) e, mais recentemente, o *National Early Warning Score 2* (NEWS 2), possibilitaram alguns passos em direção a segurança do paciente. Esses sistemas atribuem pontuações a parâmetros vitais como: (frequência respiratória, saturação de oxigênio, suplementação de O<sup>2</sup>, pressão arterial sistólica, frequência cardíaca, nível de consciência e temperatura) (CHOI et al, 2023 e CHOI et al, 2024).

A lógica desses escores é a padronização da linguagem entre as equipes de enfermagem e o corpo médico, iniciando gatilhos em direção a protocolos de resposta rápida (Times de Resposta Rápida - TRR) baseados em evidências, reduzindo a subjetividade da "preocupação clínica".

### Nuances e Contextos de Aplicações da Deterioração Clínica

A aplicação da temática varia conforme o ambiente e o perfil do paciente:

- **Contexto do Paciente Pediátrico:** Utiliza-se o PEWS (*Pediatric Early Warning Score*), focando em parâmetros ajustados pela idade e sinais comportamentais.
- **Contexto do Paciente Adulto:** Utiliza-se o MEWS (*Modified Early Warning Score*), focado em sinais vitais.
- **Contexto Obstétrico:** O MEOWS (*Modified Early Obstetric Early Warning Score*) considera as mudanças fisiológicas normais da gestação.
- **Cuidado Centrado na Família:** O estado da arte inclui a percepção da família e do paciente como um "sinal vital" adicional, reconhecendo que acompanhantes frequentemente detectam mudanças sutis de comportamento antes dos monitores. (HAMLIN et al, 2023, KIA A. et al, 2020, LOURENÇO et al, 2023).

### O Estado da Arte: Desde a Inteligência Artificial ao Monitoramento Contínuo

O horizonte atual da deterioração clínica transcende as medições intermitentes e manuais.

Estamos na era da **Vigilância Digital Contínua:**

**Algoritmos de Machine Learning:** Modelos preditivos integrados ao Prontuário Eletrônico do Paciente (PEP) analisam milhares de variáveis simultaneamente (exames laboratoriais, histórico médico e sinais vitais em tempo real). Estes modelos superam os escores tradicionais (como o MEWS) em sensibilidade e especificidade, reduzindo a "fadiga de alarmes". (PADILLA R et al, 2018, CHOI et al, 2023, e CHOI et al, 2024).

- **Sensores Wearables:** O uso de dispositivos sem fio permite a monitoração contínua de pacientes em enfermarias de baixa complexidade, identificando tendências de queda antes mesmo de um ponto de corte ser atingido.
- **Análise de Tendências (Trends):** O foco deslocou-se de um valor isolado para a trajetória do paciente. Uma frequência cardíaca que sobe consistentemente de 70 para 95 bpm em 4 horas é mais alarmante do que um valor estático de 100 bpm. (SHAMOUT et al, 2020).

A deterioração clínica laboratorial de pacientes em uma unidade de terapia intensiva (UTI) pode ser compreendida como um processo dinâmico de agravamento fisiológico com progressão de disfunção orgânica, frequentemente mensurável ao longo do tempo por trajetórias de escores, ou indicadores, de disfunção e/ou por necessidade crescente de suporte avançado (VICENT et al. 1996 e AKEL et al, 2025).

Mais particularmente, pode-se afirmar que, a deterioração clínica laboratorial na UTI raramente é um evento súbito; uma vez que, ela é, na verdade, um processo evolutivo de instabilidade fisiológica. Assim, a deterioração clínica laboratorial do paciente crítico em uma UTI manifesta-se como um declínio progressivo dos parâmetros fisiológicos, frequentemente precedendo eventos adversos graves em horas.

A deterioração clínica deixou de ser vista como um evento inevitável para ser compreendida como uma falha do sistema de vigilância. O estado da arte combina a semiótica clínica clássica com a capacidade analítica da inteligência artificial, visando o "erro zero" na detecção da instabilidade. A integração entre tecnologia de ponta e uma cultura organizacional de resposta rápida é o que define, hoje, a excelência na segurança do paciente.

**Figura 1 – Deterioração fisiológica de um paciente ao longo de 8 horas culminando em parada cardíaca e morte**



Fonte: Adaptação de Taenzer et al. (2018)

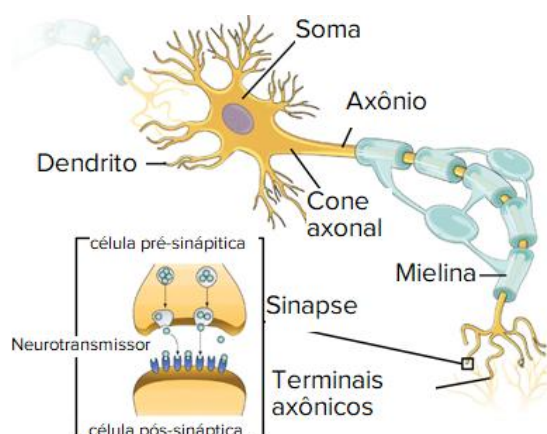
Na **figura 1**, apresenta-se a ilustração para facilitar a compreensão do início dos sinais e sintomas com a progressão das deteriorações clínicas, sejam observadas em sinais vitais e nos exames laboratoriais até o risco em potencial de morte, quando não acontece uma intervenção pontual da equipe de saúde.

## 2.2 FUNDAMENTOS DAS REDES NEURAIS ARTIFICIAIS

**KOWSHER et al. (2021)** afirmam que as redes neurais artificiais são algoritmos criados explicitamente como inspiração em redes neurais biológicas. A base das redes neurais são neurônios interconectados de acordo com o tipo de rede. Inicialmente, a ideia era criar um sistema artificial que funcionasse exatamente como o cérebro humano. (**ALAM et al., 2020**).

Na exibição da **figura 2**, observa-se a ilustração dos neurônios biológicos, embora quando comparamos com o neurônio artificial na **figura 3**, será possível compreender que apresentam objetivos distintos, porém relacionados, no contexto de suas respectivas funções dentro de sistemas complexos. O neurônio biológico, presente no sistema nervoso, tem como principal objetivo captar, processar e transmitir informações sensoriais, motoras e de integração, promovendo a coordenação de respostas adaptativas e a manutenção da homeostase do organismo. Esses neurônios operam por meio de sinais elétricos e químicos, possibilitando a comunicação eficiente entre diferentes regiões do corpo e do cérebro, sustentando funções cognitivas, motoras e sensoriais, essenciais à sobrevivência e à adaptação.

**Figura 2 – Estrutura de um Neurônio Biológico**



Fonte: Adaptação de Le et al. (2020)

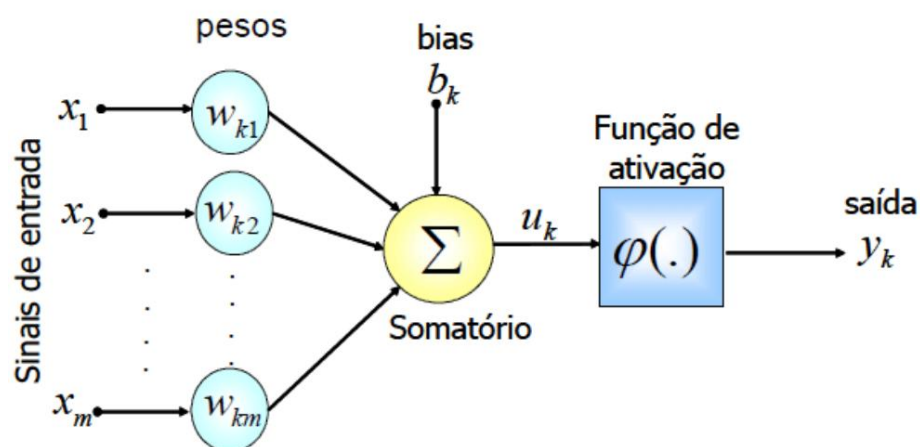
Já o neurônio artificial exibido na **figura 2**, desenvolvido a partir de modelos matemáticos e algoritmos computacionais, busca simular, de forma simplificada, as funções de um neurônio biológico com o objetivo de aprender, reconhecer padrões e resolver problemas específicos. Ele é parte integrante de redes neurais artificiais, cujo propósito é adaptar-se a diferentes inputs, ajustar seus parâmetros e gerar respostas precisas para tarefas como classificação, reconhecimento de voz, imagens ou previsão de dados. Assim, enquanto o neurônio biológico atua na sustentação e no funcionamento do organismo, o neurônio artificial tem por objetivo realizar tarefas específicas de processamento de informações de modo automatizado, contribuindo para o avanço da inteligência artificial e das tecnologias de aprendizado de máquina.

Existem muitos tipos de redes neurais, mas elas se dividem em três classes principais:

- Redes neurais *feedforward* (Unidirecional);
- Redes neurais convolucionais;
- Redes Neurais recorrentes;

Na maioria das vezes, a diferença entre elas está no tipo de neurônios que os formam e como a informação flui pela rede.

**Figura 3 – Estrutura de um Neurônio Artificial**



Fonte: Adaptado de Haykin (2009)

Na **figura 3**, é possível observar que a estrutura de um neurônio artificial e seus principais componentes. Tradicionalmente conhecido como unidade simples de processamento, o

neurônio artificial é construído por um vetor de entrada  $X [x_1, x_2, x_3, \dots x_m]$ , um vetor de pesos  $W [k_1, k_2, k_3, \dots k_m]$ , o somatório  $\sum$ , a função de ativação  $\varphi (\cdot)$  e a saída  $\gamma_k$ .

No neurônio artificial, podemos observar os dados dos vetores de entrada e pesos são direcionados para a somatória. Portanto, os dados do vetor de entrada são fornecidos externamente, os dados do vetor de pesos são iniciados randomizados. Em seguida, a função de ativação transmite ou bloqueia os dados processados no somatório, similar ao neurônio biológico. Assim, o vetor de pesos tem seus dados atualizados baseados em seu antigo valor, conforme é demonstrado a seguir na **equação 1**.

#### Equação 1: Vetor de cálculo de pesos das variáveis

$$w_i^{t+1} = w_i^t + \Delta w_i^t$$

(1)

A atualização dos dados do vetor de pesos depende da RNA utilizada, mas costuma se usar na minimização do erro entre os valores previstos pela RNA e as saídas  $\gamma_k$  desejadas. A atualização dos pesos é apresentada na equação da **equação 2** a seguir.

#### Equação 2: Atualização de cálculo dos pesos

$$\varepsilon_i = \sum w_i x_i - y_i$$

(2)

### 2.3 REGISTROS ELETRÔNICOS DE SAÚDE (RES)

Os registros eletrônicos de saúde (RES) representam uma fonte valiosa de informações clínicas, englobando dados estruturados e não estruturados ao longo do tempo dos pacientes (HIMSS, 2020). Estes dados permitem análise preditiva avançada, porém, sua heterogeneidade requer o uso de metodologias específicas de processamento para extrair informações relevantes (NGUYEN et al., 2019). Estudos demonstram que o uso eficiente dos RES pode melhorar significativamente a detecção precoce de risco de morte e outros desfechos adversos (SUNDERMANN et al. 2019).

## 2.4 MODELOS PREDITIVOS EM SAÚDE

Tradicionalmente, modelos estatísticos como regressão logística e têm sido utilizados em predição clínica (DATTA, 2022). Contudo, esses modelos apresentam limitações na captura de relações complexas e interações não lineares entre variáveis de alta dimensionalidade. Nas últimas décadas, o avanço no aprendizado de máquina (ML) e *deep learning* (DL) possibilitou a criação de modelos mais precisos, capazes de lidar com grande volume de dados e diversidade de fontes (MIOTTO et al., 2016; RAJKOMAR et al., 2019).

Modelos de ML aplicados à saúde, incluem redes neurais profundas, árvores de decisão, florestas aleatórias, e técnicas de *ensemble*, oferecendo maior acurácia, especialmente quando integrados com dados não estruturados (LIPTON Z et al., 2015). Entretanto, a interpretabilidade desses modelos ainda representam um desafio importante para sua aplicação clínica rotineira.

Modelos tradicionais de predição, como regressão logística, oferecem simplicidade na interpretação, porém enfrentam limitações ao lidar com dados altamente dimensionais e complexos (TIBSHIRANI R. 1996). Com o avanço do aprendizado de máquina, surgiram algoritmos como árvores de decisão, florestas aleatórias e redes neurais profundas, que melhor capturam relações não lineares nos dados clínicos. (OBERMEYER et al., 2016; RAJKOMAR et al., 2019)

No contexto hospitalar, modelos preditivos têm sido utilizados para classificar risco de mortalidade, readmissões, complicações cirúrgicas, entre outros desfechos clínicos (BRENNER Z et al., 2010). Sua implementação demanda atenção à interoperabilidade e à validação externa para garantir sua aplicabilidade clínica.

O PLN (Processamento de Linguagem Natural) permite a extração automática de informações de textos clínicos, que representam uma parcela relevante dos RES (CHOI A et al., 2023). Técnicas tradicionais, como análise de sentimento, extração de entidades nomeadas e análise de tópicos, evoluíram para modelos de última geração baseados em redes neurais profundas, como os *Transformers*. Estes modelos, possibilitam a compreensão semântica profunda dos textos, facilitando a incorporação de dados não estruturados na construção de modelos preditivos (GE W et al., 2022). A interpretação contextualizada de notas clínicas e laudos radiológicos melhora a acurácia do sistema de previsão.

AHMAD et al. (2023) entendem que as tecnologias de PNL podem ser usadas para uma variedade de aplicações da ciência da computação, incluindo sumarização automática, resposta a perguntas, classificação de texto, reconhecimento de entidades nomeadas, análise de

sentimentos e muito mais. Ao compreender a estrutura e o significado da linguagem humana, *insights* poderosos podem ser obtidos a partir de grandes volumes de entrada de dados formulados. (SUN et al., 2018)

Isso permitiu que as empresas obtivessem melhor acesso ao feedback do cliente e conduzissem suas tomadas de decisão com maior precisão. A PNL está remodelando a forma como as pessoas interagem com a tecnologia e os recursos de dados, permitindo que as máquinas entendam nossa linguagem natural e respondam de acordo. (FRAILE et al., 2023, KOLECK et al., 2019)

A adoção do processamento de linguagem natural na área da saúde está aumentando devido ao seu potencial reconhecido pelos sistemas de saúde para pesquisar, analisar e interpretar enormes quantidades de conjuntos de dados de pacientes. A tecnologia de PNL tem o potencial de aproveitar *insights* e conceitos relevantes de anotações clínicas que antes eram considerados pelo setor de saúde como ocultos em dados textuais. (FRAILE et al., 2023)

Os profissionais da saúde dedicam muito tempo ao registrar anotações em prontuários. Essas anotações não são facilmente extraíveis de forma que os dados possam ser analisados por um computador. (HASHIR et al., 2020) Enormes volumes de dados não estruturados de pacientes são inseridos em prontuários eletrônicos de saúde diariamente, mas é difícil para um computador ajudar os médicos e enfermeiros a agregarem esses dados críticos. A análise de grandes volumes de dados, mostra que até 80% das documentações de saúde são compostas por dados não estruturados e, portanto, em grande parte não utilizados pelos sistemas de saúde, uma vez que a ciência de dados para mineração e extração dessas informações é oneroso e consome muitos recursos. (BOULITSAKIS et al., 2023) Sem o processamento de linguagem natural na área da saúde, esses dados não estruturados não estão em um formato utilizável para que algoritmos modernos baseados em inteligência artificial os extraiam e usem de forma compreensível a crítica humana. (ASSELBERGS et al., 2019 e FANELLI et al., 2023)

Segundo BOSE et al. (2021), textos clínicos possuem informações críticas e sensíveis, muitas vezes não codificadas de forma estruturada, como diagnósticos, observações clínicas, comentários de profissionais, entre outros. A extração automática de entidades nomeadas (NER) permite converter esses textos em dados estruturados úteis para integração com sinais vitais e exames laboratoriais. (HUANG et al., 2020)

## 2.5 DESENVOLVIMENTO TÉCNICO DOS ALGORITMOS

A coleta e organização dos dados compreende sinais vitais, exames laboratoriais, provenientes do RES. Os RES representam uma base robusta de dados clínicos, consolidando informações estruturadas — sinais vitais, exames laboratoriais e não estruturadas, incluindo notas de enfermagem, notas de médicos, laudos radiológicos e textos livres (HIMSS, 2020).

A análise integrada destes dados permite a construção de modelos preditivos capazes de antever eventos adversos, contribuindo para uma assistência mais proativa (JOHNSON et al., 2020). Contudo, a heterogeneidade e complexidade desses dados demandam algoritmos de processamento capazes de extrair informações relevantes de formatos diversos (SHICKEL B et al., 2018).

## 2.6 TÉCNICAS DE RECONHECIMENTO DE ENTIDADES NOMEADAS EM SAÚDE

Um Registro Eletrônico de Saúde (RES) pode ser considerado um repositório digital de informações de saúde de um paciente. Diversas publicações trazem que em torno 80% dos dados dentro dos RES estão em um formato não estruturado, o que significa que estão contidos em documentos de texto livres codificados em linguagem humana expressiva e natural, normalmente usada para documentar procedimentos clínicos (SLATER et al., 2021). Esses dados podem ser extraídos por meio de métodos de reconhecimento de entidades nomeada conhecida como NER ou *Named Entity Recognition* ou extração de relação (RE), que são componentes cruciais do processamento de linguagem natural (PNL). (HUANG et al., 2020)

Essas tarefas envolvem identificar, extrair, associar e classificar termos clínicos, como doenças, sintomas, tratamentos, testes, medicamentos, procedimentos e partes do corpo, permitindo assim o reconhecimento de uma variedade de conceitos clínicos. A identificação de conceitos em textos na saúde é um aspecto crítico dos sistemas de suporte à decisão clínica, que são projetados para auxiliar os profissionais de saúde na tomada de decisões baseadas em dados que melhoram a qualidade dos serviços de saúde. (GIORGI et al., 2018, GIORGI et al., 2020)

Os métodos tradicionais de aprendizado de máquina abrangem redes neurais totalmente conectadas, máquinas de vetores de suporte, árvores de decisão, florestas aleatórias e outros classificadores. Esses métodos necessitam de etapas de extração de características, que normalmente são baseadas em *embeddings* de palavras. Os métodos de aprendizado profundo, por sua vez, consistem em modelos baseados em redes neurais convulsionais e recorrentes. Esses modelos, que incluem modelos de Markov Inferência Bayesiana entres outros. (VUOKKO et al., 2023)

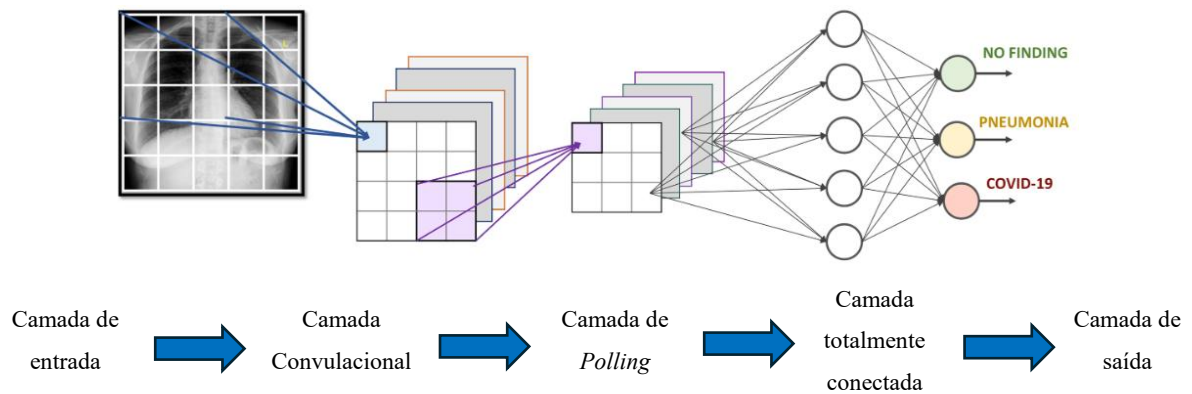
Deveremos explorar neste estudo, recursos do tipo modelos tradicionais: baseados em dicionários médicos como, ontologias e taxonomias em saúde (*UMLS, SNOMED CT, LOINC e ICD*). Para estabelecer uma associação de padrões técnicos com nosso modelo multimodal computacional.

## 2.7 REDES NEURAIAS CONVOLUCIONAIS

Elas são compostas por camadas de nós, contendo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Cada nó se conecta a outro e possui um peso e um limite associados. Se a saída de qualquer nó individual estiver acima do valor limite especificado, esse nó é ativado, enviando dados para a próxima camada da rede. Caso contrário, nenhum dado é passado para a próxima camada da rede. As redes neurais convolucionais se distinguem de outras redes neurais por seu desempenho superior com entradas de imagem, fala ou sinais de áudio. Elas possuem três tipos principais de camadas, que são: Camada convolucional; Camada de agrupamento; Camada totalmente conectada. A camada convolucional é a primeira camada de uma rede convolucional. Embora as camadas convolucionais possam ser seguidas por camadas convolucionais adicionais ou camadas de agrupamento, a camada totalmente conectada é a camada final. (WANG et al., 2019, VELDHUIS et al., 2022, ALSHWAHEEN et al., 2021).

Na **figura 4**, será possível observar que a estrutura de uma rede neural convolucional (CNN) pode ser descrita com destaque nas suas principais camadas. A rede começa com a camada de entrada, que recebe uma imagem ou um dado espacial, como um mapa de *pixels*, e o converte em uma matriz de valores numéricos. Em seguida, vem a camada convolucional, que aplica filtros (ou *kernels*) sobre a entrada, realizando operações de convolução para detectar características locais, como bordas, texturas ou padrões específicos. Esses filtros são treinados para aprender as representações mais relevantes durante o processo de treinamento. Após a camada convolucional, costuma-se usar uma camada de *pooling*, que reduz dimensionalidade e extrai as características mais importantes, além de diminuir o custo computacional.

**Figura 4 – Exemplos da estrutura de uma rede neural convolucional (CNN)**



Fonte: Adaptação de:

Convolutional neural network model based on radiological images to support COVID-19 diagnosis:  
Evaluating database biases (2025)

As operações de *pooling* podem ser *max pooling* ou *average pooling*, e ajudam a criar uma representação mais robusta e invariável a pequenas variações na entrada. Depois dessas camadas, os dados passam por uma ou mais camadas totalmente conectadas (*fully connected*), que integram as características extraídas e realizam a classificação ou regressão desejada. Essas camadas atuam como um classificador, combinando todas as informações extraídas pelas camadas anteriores para gerar uma saída final. Por fim, vem a camada de saída, que fornece o resultado final do processamento, podendo ser uma ativação *sigmoid*, *softmax* ou outra, dependendo do tipo de tarefa, como classificação binária, multiclasse ou regressão. Assim, a combinação dessas camadas permite à CNN extrair e interpretar padrões complexos em dados espaciais, tornando-a extremamente eficiente em tarefas de visão computacional, reconhecimento de imagens e outros domínios relacionados.

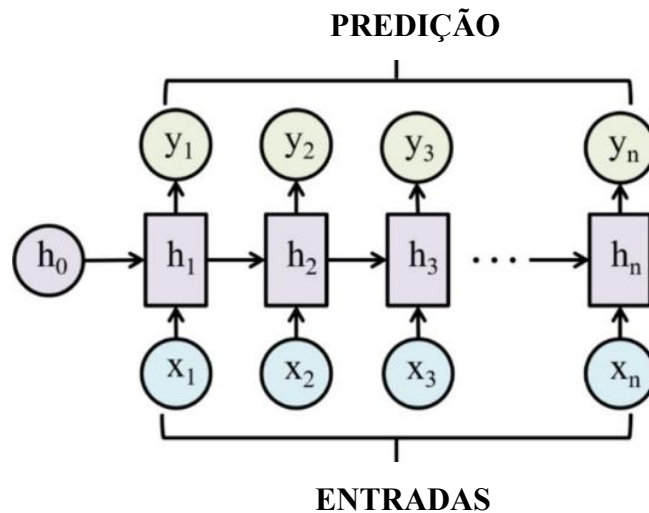
## 2.8 REDES NEURAIAS RECORRENTES

Enquanto em redes neurais *feedforward* (Unidirecionais) padrão, cada componente do vetor de entrada tem seus pesos, em redes recorrentes, cada componente compartilha os mesmos pesos. Geralmente, a vantagem das redes recorrentes é que elas compartilham pesos para cada posição do vetor de entrada. Além disso, um modelo pode processar sequências com comprimentos diferentes compartilhando os pesos. Outra vantagem é que isso reduz o número de parâmetros (pesos) que a rede precisa aprender. O princípio básico em redes recorrentes é que o vetor de entrada e algumas informações da etapa anterior (geralmente um vetor) são usados para calcular a saída e as informações passadas para a próxima etapa. Em geral, as

fórmulas usadas para calcular os valores de saída em cada etapa são chamadas de unidades (blocos). (WANG et al., 2019, VELDHUIS et al., 2022, ALSHWAHEEN et al., 2021)

Na **figura 5**, onde observamos o *Input Gate* ou (portas de entrada): A adição de informações úteis ao estado da célula é feita pelo *input gate*. Primeiro, a informação é regulada usando a função sigmoide que filtra os valores a serem lembrados de forma similar ao *forget gate* usando as entradas  $h_{t-1}$  e  $x_t$ . Então, um vetor é criado usando a função tanh que dá saída de -1 a +1, que contém todos os valores possíveis de  $h_{t-1}$  e  $x_t$ . Os valores do vetor e os valores regulados são multiplicados para obter as informações úteis.

**Figura 5 – Estrutura de uma rede neural recorrente**



Fonte: Adaptado de Haykin (2009)

## 2.9 REDES NEURAIAS RECORRENTES LSTM

É um subtipo especial de rede neural recorrente. Especificamente, essa arquitetura foi introduzida para resolver problemas de gradientes que desaparecem e explodem. Além disso, esse tipo de rede é mais adequado para manter conexões de longo alcance, reconhecendo a relação entre os valores no início e no fim de uma sequência, segundo **CHOI et al. (2024)**. O modelo LSTM introduz expressões, em particular, como portas.

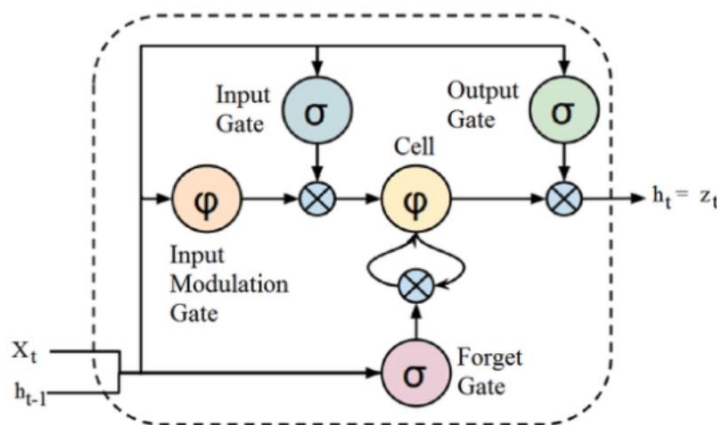
**Existem três tipos de portas:**

- Porta de esquecimento *Forget Gate* – controla quanta informação a célula de memória receberá da célula de memória da etapa anterior;

- Porta de atualização **Input Gate** (entrada) – decide se a célula de memória será atualizada. Além disso, controla a quantidade de informações que a célula de memória atual receberá de uma célula de memória potencialmente nova;
- Porta de saída **Output Gate** – controla o valor do próximo estado oculto;

Na **figura 6**, onde observamos o **Forget Gate** ou (portas de esquecimento): As informações que não são mais úteis no estado da célula e são removidas. Duas entradas:  $x_t$  (entrada no momento específico) e  $h_{t-1}$  (saída de célula anterior) são alimentadas ao *gate* (porta) e multiplicadas por matrizes de peso, seguidas pela adição do bias. O resultante é passado por uma função de ativação que fornece uma saída binária.

**Figura 6 – Arquitetura de uma rede neural LSTM**



Fonte: Adaptado de Haykin (2009)

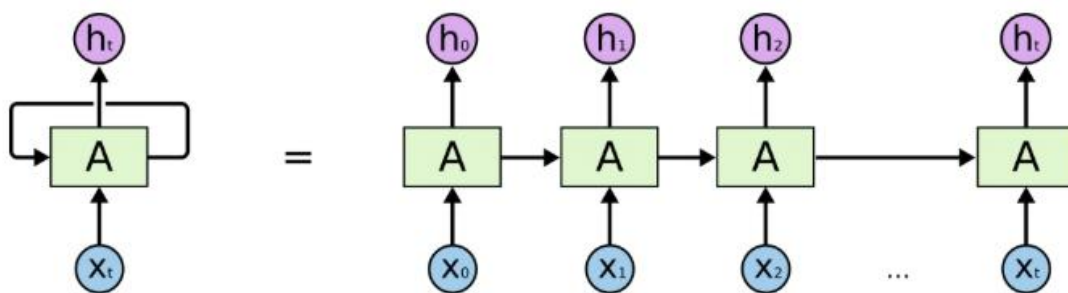
Se para um determinado estado de célula a saída for 0, a informação é esquecida e para a saída 1, a informação é retida para uso futuro.

Também na **figura 5**, observa-se o **Output Gate** ou (portas de saída): A tarefa de extrair informações úteis do estado da célula atual para ser apresentadas como uma saída é feita pelo *output gate*. Primeiro, um vetor é gerado aplicando a função **tanh** na célula. Então, a informação é regulada usando a função sigmóide que filtra os valores a serem lembrados usando as entradas  $h_{t-1}$  e  $x_t$ . Os valores do vetor e os valores regulados são multiplicados para serem enviados como uma saída e entrada para a próxima célula.

Na **figura 7**, visualiza-se como uma rede neural LSTM (*Long Short-Term Memory*) simples, composta por uma única camada de células RNN projetadas para lidar com sequências

de dados, permitindo a captura de dependências de longo prazo. Sua estrutura começa com uma camada de entrada, que recebe uma sequência de dados, como palavras ou valores temporais, e converte cada elemento em um vetor numérico. Essa arquitetura permite que a rede capture dependências temporais ou espaciais complexas, sendo eficaz em tarefas como previsão de séries temporais, reconhecimento de fala e processamento de linguagem natural.

**Figura 7 – Estrutura de uma rede neural LSTM**



Fonte: Adaptado de Haykin (2009)

## 2.10 REDES NEURAS RECORRENTES LSTM BIDIRECIONAIS

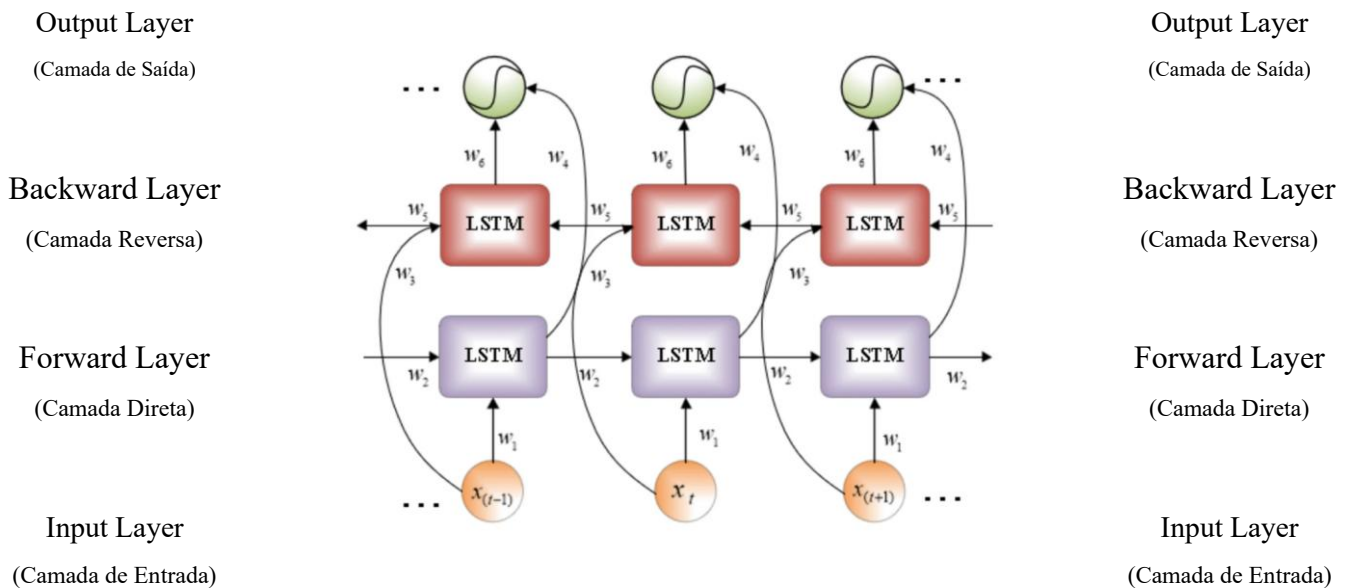
LSTM bidirecional (BiLSTM) é uma rede neural recorrente usada principalmente no processamento de linguagem natural. Ao contrário da *LSTM* padrão, a entrada flui em ambas as direções e ela é capaz de utilizar informações de ambos os lados. Também é uma ferramenta poderosa para modelar as dependências sequenciais entre palavras e frases em ambas as direções da sequência. Em resumo, o BiLSTM adiciona mais uma camada LSTM, o que inverte a direção do fluxo de informações. Resumidamente, isso significa que a sequência de entrada flui para trás na camada LSTM adicional. Em seguida, combinamos as saídas de ambas as camadas LSTM de diversas maneiras, como média, soma, multiplicação ou concatenação. (WANG et al., 2019, VELDHUIS et al., 2022, ALSHWAHEEN et al., 2021)

O aprendizado de máquina está se tornando cada vez mais avançado com a progressão das tecnologias de ponta. Como os algoritmos existentes não oferecem um desempenho de aprendizado satisfatório na maioria das vezes, é necessário continuar atualizando os algoritmos atuais incessantemente. A hibridização de dois ou mais algoritmos pode potencialmente aumentar o desempenho do modelo projetado. Embora LSTM e BiLSTM sejam dois algoritmos

excelentes e amplamente utilizados, ainda há espaço para melhorias em termos de precisão por meio do método de hibridização. Assim, as vantagens dos algoritmos RNN e ANN podem ser obtidas simultaneamente. (WANG et al., 2019, VELDHUIS et al., 2022, ALSHWAHEEN et al., 2021)

Na **figura 8**, observa-se que, a estrutura de uma rede neural BiLSTM (*Bidirectional Long Short-Term Memory*) pode ser explicada de forma técnica considerando seus principais componentes. Ela começa com uma camada de entrada (*input layer*), que recebe uma sequência de dados, como palavras ou valores temporais, e converte cada elemento em um vetor de *features* ou *embeddings*. Essa sequência de vetores é processada simultaneamente por duas camadas LSTM independentes: a camada *forward* (*forward layer*) e a camada *backward* (*backward layer*).

**Figura 8 – Conexões de uma rede neural LSTM**



Fonte: Adaptado de Haykin (2009)

A camada *forward* percorre a sequência na ordem natural, do primeiro ao último elemento, capturando o contexto passado até o momento atual. Já a camada *backward* processa a sequência na ordem reversa, do último ao primeiro elemento, permitindo que a rede também aprenda o contexto futuro. As saídas dessas duas camadas, que representam a informação adquirida de ambos os lados, são então concatenadas em cada passo para formar uma

representação mais completa da sequência. Essa representação combinada é enviada para a camada de *output* (*output layer*), que pode ser uma camada densa, responsável por interpretar a informação e realizar tarefas específicas, como classificação ou previsão. Dessa forma, a arquitetura BiLSTM integra informações de ambos os contextos temporal e sequencial, tornando-se especialmente eficaz em tarefas que requerem uma compreensão profunda do contexto completo de uma sequência.

## 2.11 REVISÃO DE ESCOPO DA LITERATURA SOBRE USO DE INTELIGÊNCIA ARTIFICIAL EM REGISTROS ELETRÔNICOS DE SAÚDE

A Revisão de Escopo é uma metodologia de pesquisa que tem ganhado crescente atenção no campo da ciência, especialmente na fase inicial de investigações e na elaboração de revisões integradas de literatura. Ela se diferencia de revisões sistemáticas e revisões narrativas ao proporcionar uma análise ampla e abrangente do estado da arte sobre um determinado tema, permitindo a identificação de lacunas, tendências e áreas emergentes.

Segundo **GRANT e BOOTH (2009)**, a Revisão de Escopo é uma técnica de síntese de evidências que mapeia a maior quantidade possível de literatura relevante para um tema específico, sem, inicialmente, estabelecer critérios tão restritivos quanto uma revisão sistemática. Sua estrutura permite que pesquisadores entendam a amplitude do conhecimento disponível, incluindo diferentes tipos de estudos, metodologias e resultados, consolidando uma visão geral do campo de estudo.

Ao contrário da revisão sistemática, que busca responder perguntas específicas e realizar uma avaliação crítica da qualidade dos estudos, a Revisão de Escopo tem como foco principal a descrição e organização do universo de produção científica, estabelecendo os limites do tema e identificando possíveis áreas para futuras investigações.

A fundamentação teórica da Revisão de Escopo apoia-se em metodologias desenvolvidas por **PETERS et al (2020)**, que propuseram etapas sistemáticas para sua realização: identificação da pergunta da pesquisa, busca de estudos, seleção, extração de dados, análise e relato dos resultados. Essas etapas garantem rigor e transparência ao processo, mesmo que o objetivo seja mapear a produção científica ampla.

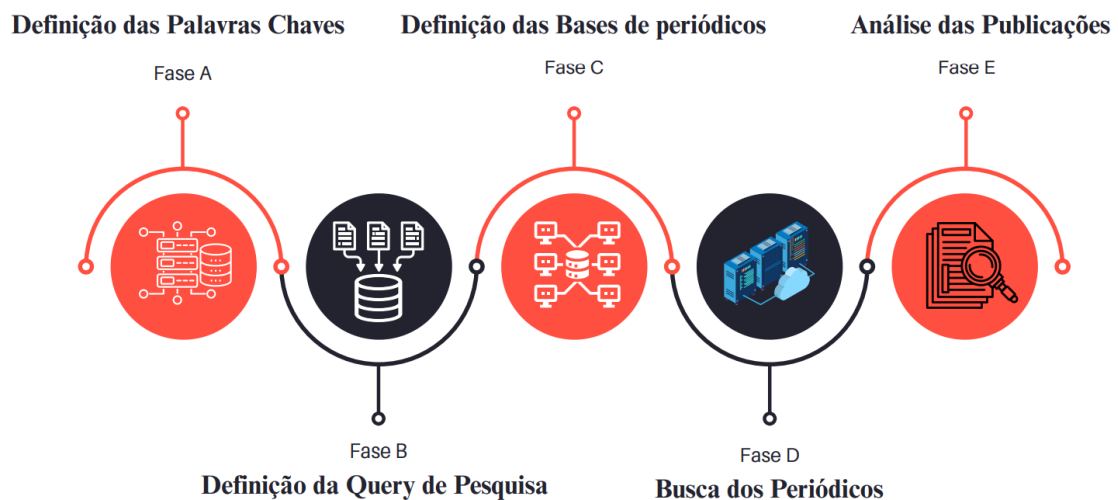
A técnica permite a inclusão de diferentes tipos de fontes, como artigos empíricos, revisões, relatórios técnicos e documentos não acadêmicos, promovendo uma compreensão multifacetada do tema. Em contextos de tese de doutorado, o uso da Revisão de Escopo pode

auxiliar na delimitação do objeto do estudo, na construção de quadros teóricos e na identificação de lacunas que justificam novas investigações mais específicas.

A revisão de escopo é particularmente relevante em contextos em que há grande volume de literatura heterogênea ou pouco consolidada. Sua aplicação em uma tese de doutorado permite ao pesquisador estabelecer um panorama crítico, oferecer uma fundação sólida para o desenvolvimento de hipóteses e teorizações, além de orientar futuras linhas de pesquisa.

Além disso, essa abordagem contribui para evitar redundâncias e promover a inovação científica ao identificar áreas pouco exploradas ou com potencial de avanço. Como ferramenta preliminar, ela também fortalece a justificativa do estudo, ao evidenciar a necessidade e a relevância do tema abordado.

**Figura 9: Organização das etapas da Revisão de Escopo de Pesquisa da Literatura**



Fonte: Autor (2025)

Na **etapa A – Definição das palavras chaves:** foram elencadas as palavras chaves relacionadas ao tema da pesquisa o uso de inteligência artificial com registros eletrônicos em saúde. Optou-se pelo uso das palavras chaves em inglês e português. Decidiu-se por esta escolha, em razão da oferta de publicações no idioma em inglês, além de uso do idioma português, pela natureza e origem da pesquisa. Palavras Chaves: (1) Deteriorações clínicas, (2) Mortalidade, (3) Registros Eletrônicos de Saúde, (4) Aprendizado de Máquina, (5) Aprendizagem de Máquina Profundo, (6) Inteligência Artificial, (7) Processamento de Linguagem Natural e (8) Paciente Adulto. MESH: *titles with keywords such as:* (1) *Clinical Deteriorations*, (2) *Mortality*, (3) *Electronic Health Records*, (4) *Machine Learning*, (5) *Deep*

*Machine Learning, (6) Artificial Intelligence, (7) Natural Language Processing, and (8) Adult Patient.*

Na **etapa B – Definição da query de pesquisa**: definiu-se as palavras chaves com base no desenho da pesquisa da seguinte forma: “Deteriorações clínicas” and “Mortalidade” and “Registros Eletrônicos de Saúde” and “Aprendizado de Máquina” and “Aprendizagem de Máquina Profundo” and “Inteligência Artificial Generativa” and “Processamento de Linguagem Natural” and “Paciente Adulto” OR “*Clinical Deteriorations*” and “*Mortality*” and “*Electronic Health Records*” and “*Machine Learning*” and “*Deep Machine Learning*” and “*Artificial Intelligence*” and “*Natural Language Processing*” and “*Adult Patient*”. A compreensão e formulação de uma *query* de pesquisa eficiente é fundamental para o desenvolvimento de uma revisão de escopo, pois ela orienta a busca por informações relevantes e confiáveis em bancos de dados acadêmicos e outras fontes de conhecimento. Uma *query* bem elaborada deve refletir claramente o problema ou o tema central da investigação, utilizando termos específicos, operadores *booleanos*, truncamentos e filtros que aumentem a precisão dos resultados obtidos. Nesse contexto, a construção da *query* exige uma compreensão profunda dos conceitos principais e das possíveis variações terminológicas, bem como uma estratégia de combinação que maximize a abrangência sem comprometer a relevância das informações recuperadas. Além disso, a revisão de escopo muitas vezes demanda uma abordagem iterativa, onde a *query* inicialmente criada é refinada com base na análise preliminar dos resultados, identificando lacunas, termos novos ou variantes que possam melhorar a busca. A importância de uma *query* da pesquisa bem estruturada reside na sua capacidade de garantir que os resultados obtidos sejam representativos do universo de estudos disponíveis, possibilitando uma análise ampla e criteriosa do tema, além de evitar a exclusão inadvertida de estudos relevantes. Dessa forma, a elaboração de uma *query* adequada é uma etapa crítica, que exige reflexão, precisão e conhecimentos sobre as bases de dados e estratégias de busca, contribuindo assim para uma fundamentação sólida na revisão de escopo e para a credibilidade do trabalho científico desenvolvido.

Na **etapa C – Definição das Bases de Periódicos**: definiu-se por uso das seguintes bases de periódicos: *Medical Literature Analysis and Retrieval System*, Online (MEDLINE), Excerpta Medica dataBASE (EMBASE), *The Cumulative Index to Nursing and Allied Health Literature* (CINAHL), via plataforma EBSCO, *Scopus* (Elsevier) e *Web of Science* e IEEE Xplore.

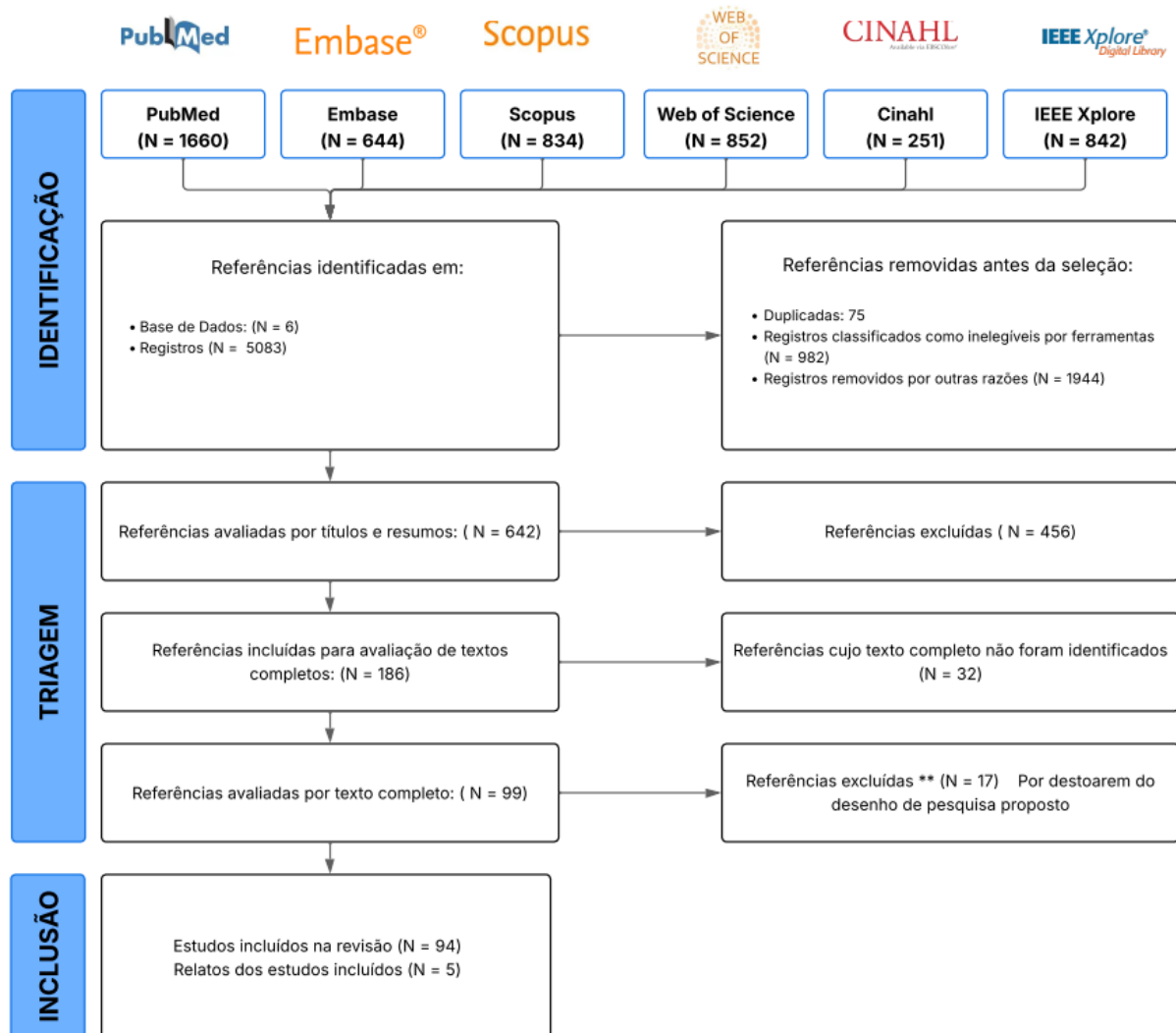
Na **etapa D – Busca dos Periódicos**: realizou-se uma busca por publicações nas bases escolhidas na fase C com o *query* de pesquisa definida na fase B. Desta forma, definiu-se que

o período temporal de busca de publicações foi de janeiro 2025 a agosto de 2025.

Na **etapa E – Análise das Publicações**: aplicação do *checklist Extension for Scoping Reviews* (*Checklist* para revisão de escopo baseado em método PRISMA, o processo sistemático é usado como uma síntese de literatura integrativa para mapear a amplitude das evidências sobre o tópico, um conceito ou uma questão específica. O PRISMA foi desenvolvido inicialmente para orientar a elaboração e avaliação de revisões sistemáticas e meta-análises, mas seus princípios podem ser adaptados para a condução de revisões de escopo, contribuindo para uma análise crítica e sistemática das publicações selecionadas. Esses critérios permitem que o pesquisador avalie aspectos essenciais, como a clareza na definição do objetivo, critérios de inclusão e exclusão, métodos de busca e seleção, além da avaliação da qualidade dos estudos. A aplicação do *checklist* de PRISMA assegura que o processo de seleção seja transparente e livre de vieses, possibilitando uma síntese mais confiável dos estudos disponíveis e facilitando a identificação de lacunas no conhecimento. Além disso, esse enfoque promove uma padronização na análise, possibilitando comparações entre diferentes revisões e aumentando a credibilidade do trabalho científico realizado. Contudo, é importante reconhecer que alguns critérios do PRISMA necessitam de ajustes ou complementações específicas para contextos de revisão de escopo, dado seu foco principal em revisões sistemáticas mais específicas. Assim, a adoção desses critérios de análise não apenas fortalece a metodologia da revisão, mas também contribui para o avanço do conhecimento, ao fornecer uma avaliação detalhada e replicável das publicações científicas relacionadas ao tema investigado.

De posse dos resultados descritos na **figura 10** nota-se que foram mapeadas 5.083 publicações encontradas nas bases de dados. Das quais, apenas 101 foram selecionadas para análise, que atenderam critérios de inclusão e exclusão. Portanto, 4.982 de 5.083 foram descartadas por destoarem do desenho de pesquisa. Desta forma, as 101 publicações escolhidas, se equivalem a 1,99% da amostra total.

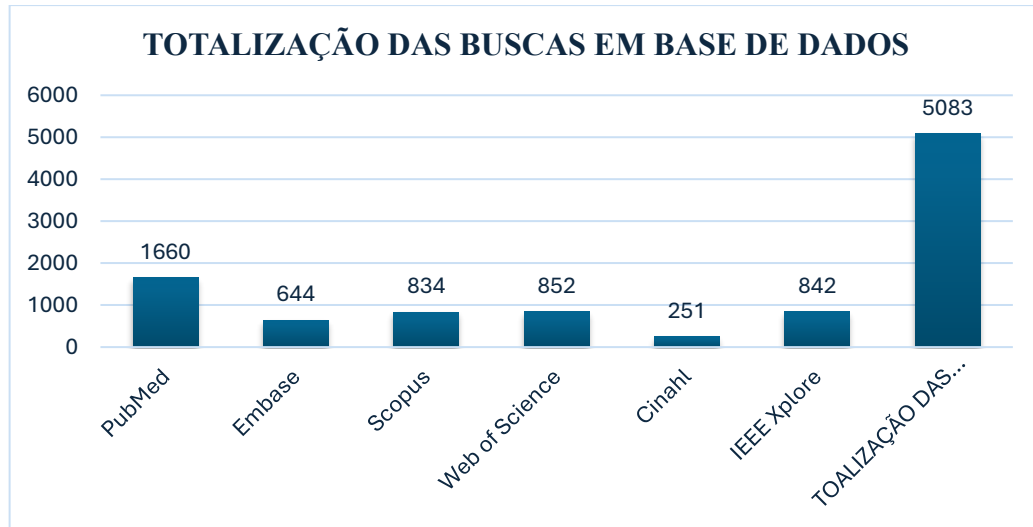
Figura 10 – Checklist de Revisão de Escopo baseado em método PRISMA



Fonte: Autor (2025)

Em observação ao **gráfico 1**, observa-se que das 5.083 (N - Total) publicações, a maioria estavam indexadas na base *PUBMED*, com 1.660 (N) sendo responsável por 32,66% do (N) total, na base *EMBASE*, com 644 (N) sendo responsável por 12,67%, na base *SCOPUS*, com 834 (N) sendo responsável por 16,41%, na base *WEB OF SCIENCE*, com 852 (N) sendo responsável por 16,76%, na base *CINAHL*, com 251 (N) sendo responsável por 4,94% e por fim na base *IEE EXPLORE – Digital Library* com 842 (N), sendo responsável por 16,57% da amostra total.

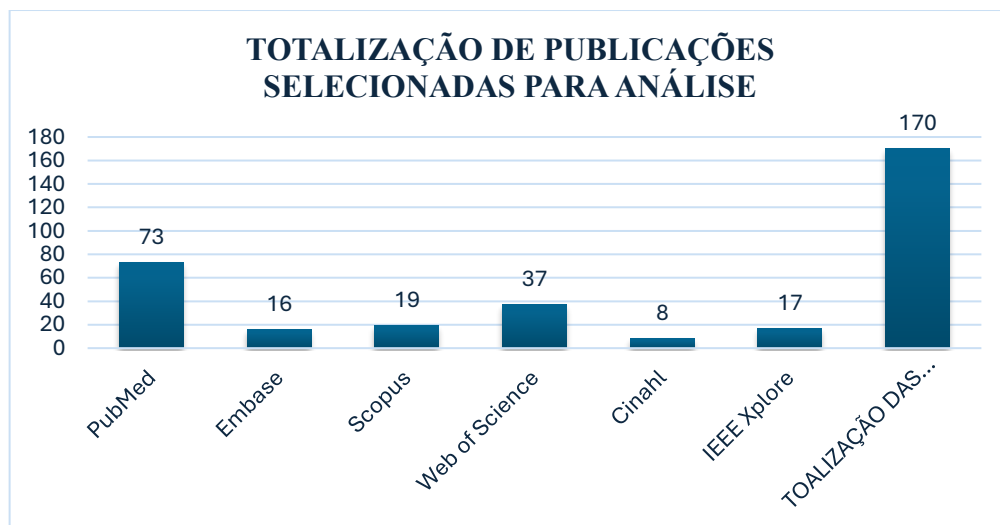
**Gráfico 1 – Totalização de Buscas em Bases de Dados**



Fonte: Autor (2025)

Em observação ao **gráfico 2**, percebe-se que das 170 (N - Total) publicações selecionadas para análise, estavam base PUBMED, com 73 (N) sendo responsável por 42,94% do (N) total, na base EMBASE, com 16 (N) sendo responsável por 9,41%, na base SCOPUS, com 19 (N) sendo responsável por 11,18%, na base WEB OF SCIENCE, com 37 (N) sendo responsável por 21,76%, na base CINAHL, com 8 (N) sendo responsável por 4,71% e por fim na base IEEEEXPLORE – *Digital Library* com 17 (N), sendo responsável por 10,0% da amostra total de selecionadas.

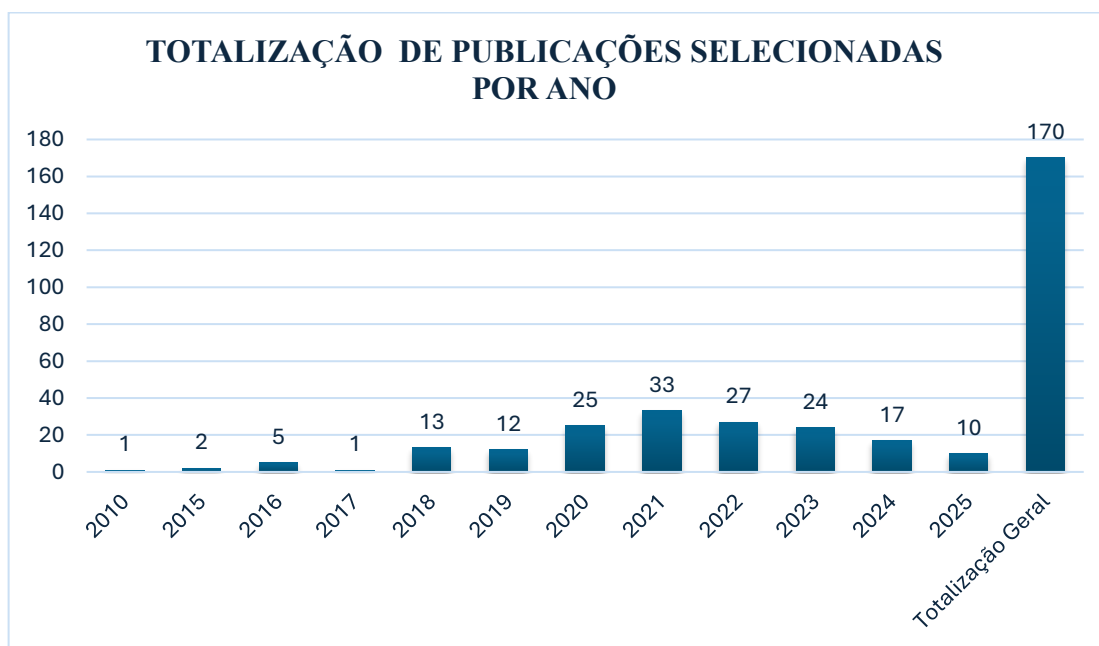
**Gráfico 2 – Totalização de Buscas em Bases Selecionadas**



Fonte: Autor (2025)

Nota-se no **gráfico 003**, que das 170 (N) publicações selecionadas para análise final, 20 destas, foram publicadas no ano de 2025 correspondente a 5,88%, 17 foram publicadas no ano de 2024 correspondente a 10,%, 24 foram publicadas no ano de 2023 correspondente a 14,12%, 27 foram publicadas no ano de 2022 correspondente a 15,88%, 33 foram publicadas no ano de 2021 correspondente a 19,41%, 25 foram publicadas no ano de 2020 correspondente a 14,71%, 12 foram publicadas no ano de 2019 correspondente a 7,06%, 13 foram publicadas no ano de 2018 correspondente a 7,65%, 1 foi publicada no ano de 2017 correspondente a 0,59%, 5 foram publicadas no ano de 2016 correspondente a 2,94%, 2 foram publicadas no ano de 2015 correspondente a 1,18% e por fim, 1 única publicação em 2010, embora > 10 anos, mas com relevância científica correspondente a 0,59%.

**Gráfico 3 – Totalização de Publicações Selecionadas por Ano**

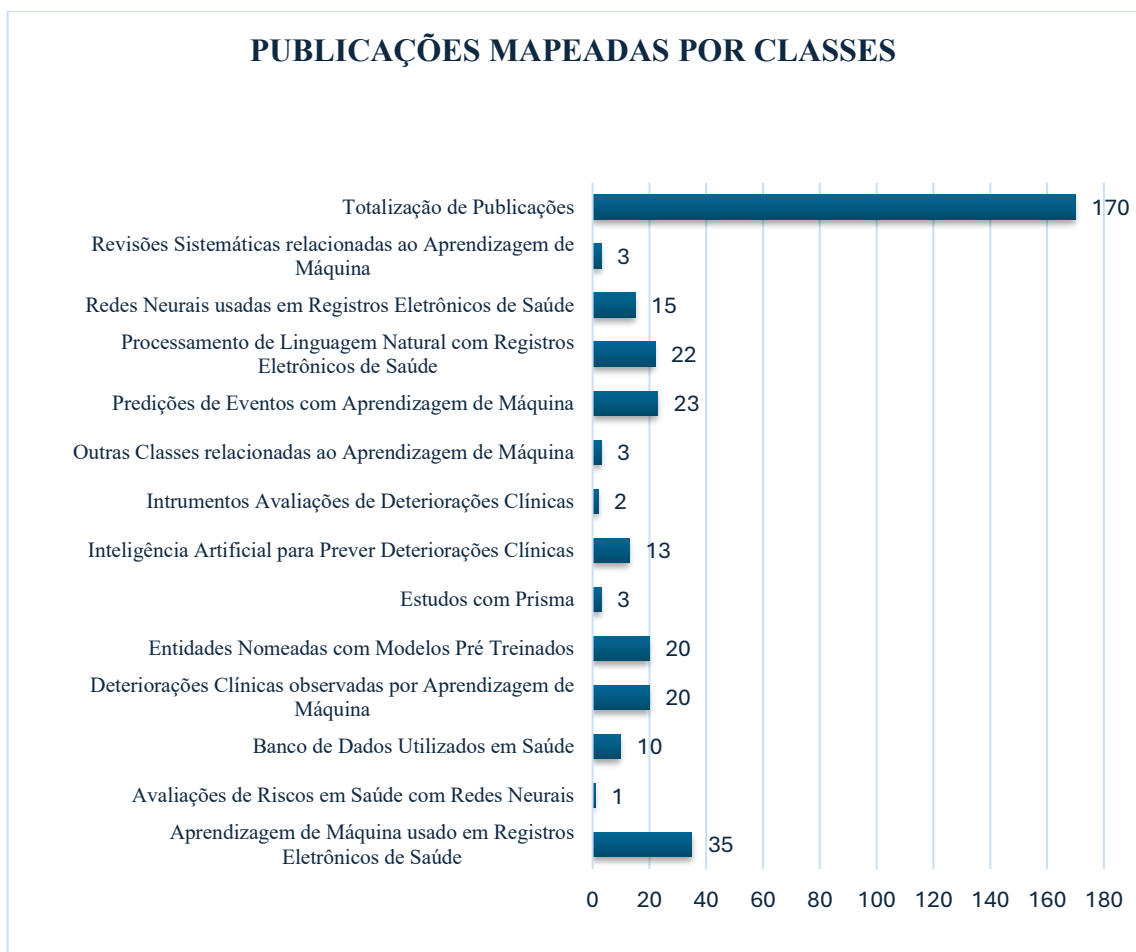


Fonte: Autor (2025)

Com base no volume de publicações selecionadas, entendeu-se relevante classificar as publicações em classes, a fim de categorizar as avaliações e nortear os achados de cada estudo observado. Desta forma, observamos no **gráfico 4**, onde na **Classe: Aprendizagem de máquina usados em registros eletrônicos de saúde**, com 35 publicações, correspondente a 20,59%; na **Classe: Avaliações de riscos em saúde com redes neurais** com 1 publicação, correspondente a 0,59%; na **Classe: Banco de dados utilizados em saúde**, com 10 publicações, correspondente a 5,88%; na **Classe: Deteriorações clínicas observadas por aprendizagem de máquina** com 20 publicações, correspondente a 11,76%; na **Classe: Entidades nomeadas**

com modelos pré treinados, com 20 publicações, correspondente a 11,76%; na **Classe: Estudos com metodologia Prisma**, com 3 publicações, correspondente a 1,76%; na **Classe: Inteligência artificial para prever deteriorações clínicas**, com 13 publicações, correspondente a 7,65%; na **Classe: Instrumentos de avaliações de deteriorações clínicas**, com 2 publicações, correspondente a 1,18%; na **Classe: Predições de eventos com aprendizagem de máquina**, com 23 publicações, correspondente a 13,53%; na **Classe: Processamento de linguagem natural com registros eletrônicos de saúde**, com 22 publicações, correspondente a 12,94%; na **Classe: Redes neurais usadas em registros eletrônicos de saúde**, com 15 publicações, correspondente a 8,82%; na **Classe: Outras classes relacionadas ao aprendizagem de máquina**, com 3 publicações, correspondente a 1,76% e por fim na **Classe: Revisões sistemáticas relacionadas ao aprendizagem de máquina**, com 3 publicações, correspondente a 1,76%.

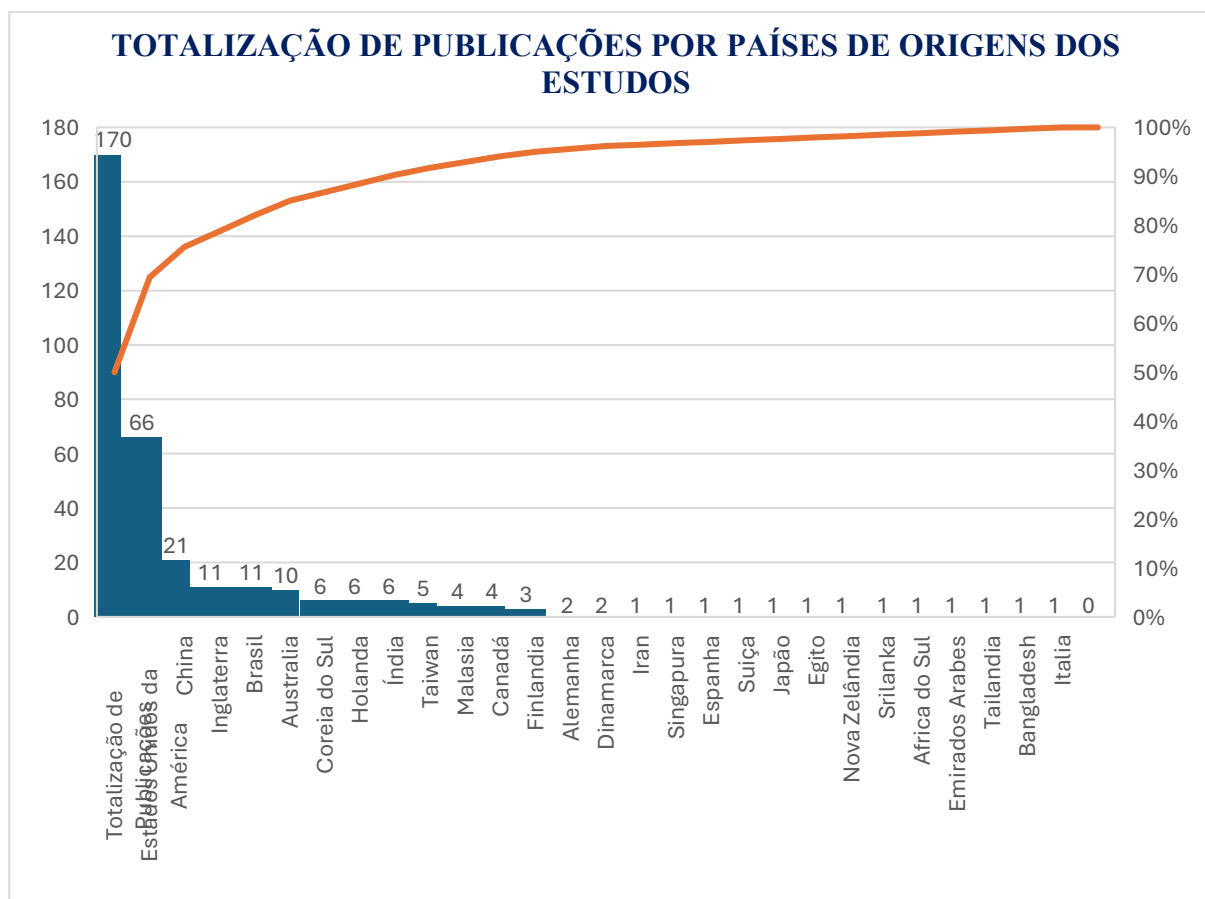
**Gráfico 4 – Totalização de Publicações Mapeadas por Classes**



Fonte: Autor (2025)

Observa-se no **gráfico 5**, que das 170 (N) do total de publicações, com 66 *papers*, os Estados Unidos da América que lideram o *ranking*, que corresponde a 38,82%, seguido pela China com 21 publicações, que corresponde a 12,35%, com 11 publicações, a Inglaterra e Brasil, que corresponde a 6,47%, com 10 publicações, a Austrália, que corresponde a 5,88%, com 6 publicações Coreia do Sul, Índia e Holanda, que corresponde a 3,53%, com 5 publicações Taiwan, que corresponde a 2,94%, com 4 publicações Malásia e Canadá, que corresponde a 2,35%, com 3 publicações a Finlândia, que corresponde a 1,76%, com 2 publicações a Alemanha e Dinamarca, que corresponde a 1,18%, e por fim com 1 publicação cada, estão iran, Singapura, Espanha, Suíça, Japão, Egito, Nova Zelândia, Sirilanka, África do Sul, Emirados Árabes, Tailândia, Blangadesh e Itália, que corresponde a 0,59%.

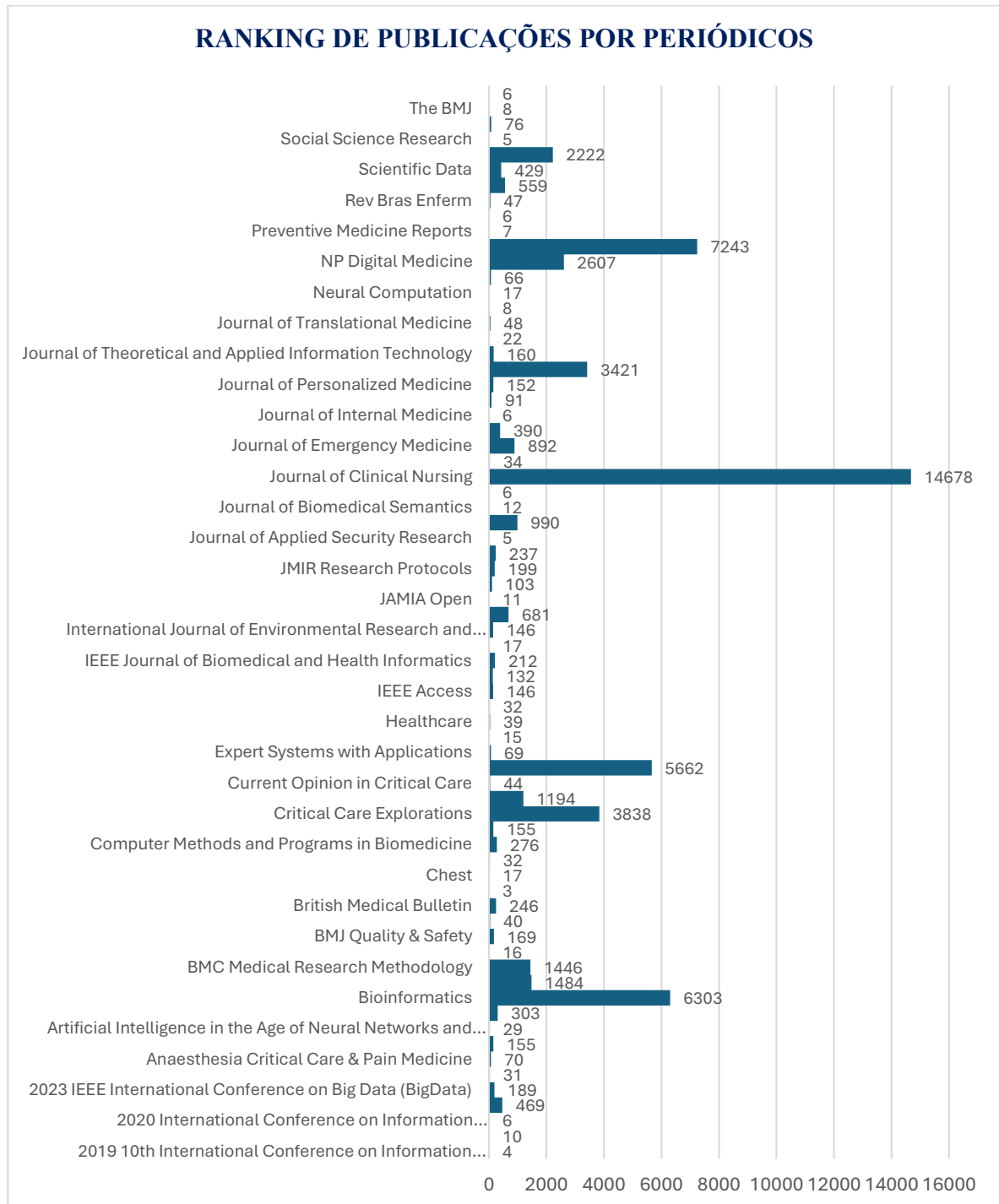
**Gráfico 5 – Totalização de Publicações Mapeadas por Países de Origens**



Fonte: Autor (2025)

Com base no **Gráfico 6**, segundo número de citações, mapeou-se nos *Journals* selecionados para seus respectivos números. Destacamos os 5 melhores ranqueados, segundo número de citações: Em 1º Lugar: *Journal of Clinical Nursing* com 14.678 citações (Segundo *SCIMAGO* – (Q1) - SJR 2024 1.322 e H-Index 131). Em 2º Lugar: *PLoS ONE* com 7.243 citações (Segundo *SCIMAGO* – (Q1) - SJR 2024 0.803 e H-Index 467). Em 3º Lugar: *Bioinformatics* com 6.303 citações (Segundo *SCIMAGO* – (Q1) - SJR 2024 2.451 e H-Index 486). Em 4º Lugar: *European Heart Journal* com 5.662 citações (Segundo *SCIMAGO* – (Q1) - SJR 2024 4.987 e H-Index 367). E em 5º Lugar: *Critical Care Explorations* com 3.838 citações (Segundo *SCIMAGO* – (Q1) - SJR 2024 0.855 e H-Index 34).

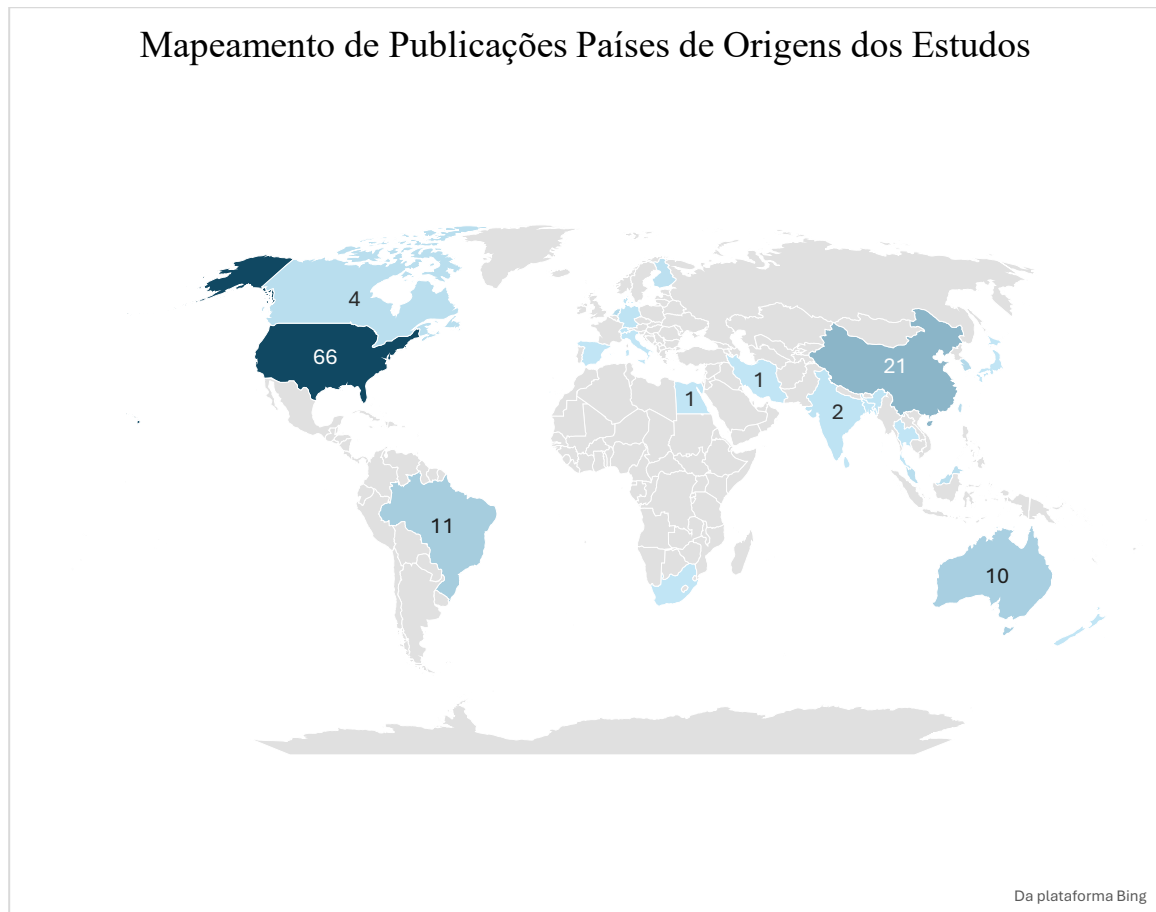
**Gráfico 6 – Ranking dos periódicos mais bem classificados, segundo as citações**



Fonte: Autor (2025)

No **gráfico 7**, observa-se a distribuição dos continentes e países onde concentra-se o maior número de publicações selecionadas neste estudo.

**Gráfico 7 – Geo-Mapeamento das Publicações Países de Origens dos Estudos**



Fonte: Autor (2025)

A implementação da aprendizagem de máquina (AM) em registros eletrônicos de saúde (RES) tem surgido como uma das áreas mais promissoras na interface entre tecnologia e assistência em saúde. Com o crescimento exponencial na geração de dados clínicos digitais, a habilidade de extrair informações relevantes de grandes volumes de dados tornou-se essencial para aprimorar a eficiência, a precisão diagnóstica e os desfechos clínicos.

A AM fornece ferramentas capazes de identificar padrões complexos que frequentemente escapam à observação humana, permitindo a previsão de eventos adversos, a classificação automática de diagnósticos e a personalização de intervenções, contribuindo para um cuidado mais preciso e personalizado. No entanto, apesar do potencial, o uso dessas técnicas

ainda encontra obstáculos relevantes, como a qualidade variada dos dados, preocupações com a privacidade e a segurança, além da necessidade de modelos interpretáveis que possam ser compreendidos pelos profissionais de saúde. Ademais, questões relacionadas à interoperabilidade entre diferentes sistemas de RES dificultam a implementação homogênea dessas soluções. Percebe-se, portanto, uma trajetória de crescimento e inovação, porém permeada por dificuldades que exigem reflexão crítica e esforços colaborativos entre desenvolvedores, profissionais de saúde e demais partes interessadas.

A previsão futura indica uma evolução contínua das metodologias de AA, com avanços em algoritmos mais interpretáveis e na integração de tecnologias como a inteligência artificial explicável, que deverão facilitar a adoção plena em ambientes clínicos. Assim, compreender as aplicações atuais, os obstáculos enfrentados e as tendências emergentes no uso de AA em RES é fundamental para orientar uma revisão de escopo que possa contribuir para o desenvolvimento de estratégias mais seguras, eficazes e éticas na implementação dessas ferramentas no âmbito da saúde digital.

A seguir no quadro 01, veremos um resumo de publicações analisadas no estudo sobre Aprendizagem de máquina usados em registros eletrônicos de saúde.

**Quadro 1 – Publicações que abordaram: Aprendizagem de Máquina em Registros Eletrônicos de Saúde**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Scalable and accurate deep learning with electronic health records	Estados Unidos da América	2227	RAJKOMAR et al. 2018
Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records	Estados Unidos da América	1359	MIOTTO et al. 2016
Development and evaluation of machine learning training strategies for neonatal mortality prediction using multicountry data	Brasil	0	SILVA et al. 2025
Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados	Brasil	0	OLIVEIRA et al. 2025
Autonomous International Classification of Diseases Coding Using Pretrained Language Models and Advanced Prompt Learning Techniques: Evaluation of an Automated Analysis System Using Medical Text	China	2	ZHUANG et al. 2025
A Survey of Large Language Models for Healthcare: from Data, Technology, and Applications to Accountability and Ethics	Singapura	177	HE et al. 2023

Benchmarking deep learning models on large healthcare MIMIC datasets	China	332	Purushotham et al. 2018
Predicting the Mortality of ICU Patients by Topic Model with Machine-Learning Techniques	Taiwan	9	CHIU et al. 2022
A deep learning approach for inpatient length of stay and mortality prediction	Estados Unidos da América	6	CHEN, J. et al. 2020
Integrating Structured and Unstructured EHR Data for Predicting Mortality by Machine Learning and Latent Dirichlet Allocation Method	Taiwan	21	CHIU et al. 2023
Unstructured clinical notes within the 24 hours since admission predict short, mid & long-term mortality in adult ICU patients	Estados Unidos da América	33	MAHBUB et al. 2022
Machine learning: An applied econometric approach	Estados Unidos da América	1446	MULLAINATHAN et al. 2017
Knowledge Discovery: Methods from data mining and machine learning	Estados Unidos da América	145	SHU,et al. 2023
The Impact of a Machine Learning Early Warning Score on Hospital Mortality: A Multicenter Clinical Intervention Trial	Estados Unidos da América	38	WINSLOW et al. 2022
Development of a machine learning-based clinical decision support system to predict clinical deterioration in patients visiting the emergency department	Coreia do Sul	22	CHOI et al. 2023
Machine learning for real-time prediction of complications in critical care: a retrospective study	Alemanha	268	MEYER et al. 2018
Less is more: Detecting clinical deterioration in the hospital with machine learning using only age, heart rate, and respiratory rate	Estados Unidos da América	33	AKEL et al. 2021
Machine learning model predicts short-term mortality among prehospital patients: A prospective development study from Finland	Finlândia	5	TAMMINEN et al. 2021
Using machine learning to improve the accuracy of patient deterioration predictions: Mayo Clinic Early Warning Score (MC-EWS)	Estados Unidos da América	39	ROMERO-BRUFU et al. 2021
Machine learning-based 28-day mortality prediction model for elderly neurocritically ill patients	China	2	YUAN J et al. 2025
Interpretable machine learning model for predicting the prognosis of antibody positive autoimmune encephalitis patients	China	1	GUO et al. 2025
Development of a machine learning model for the prediction of the short-term mortality in patients in the intensive care unit	China	6	YANG et al. 2022
A review of deep learning models and online healthcare databases for electronic health records and their use for health prediction	Emirados Árabes	6	NASARUDIN et al. 2024
Evaluation of machine learning-based models for prediction of clinical deterioration: A systematic literature review	Australia	22	JAHANDIDEH et al. 2023
Vital Signs Data and Probability of Hospitalization, Transfer to Another Facility, or Emergency Department Death Among Adults Presenting for Medical Illnesses to the Emergency Department at a Large Urban Hospital in the United States	Estados Unidos da América	7	ASIIMWE et al. 2020

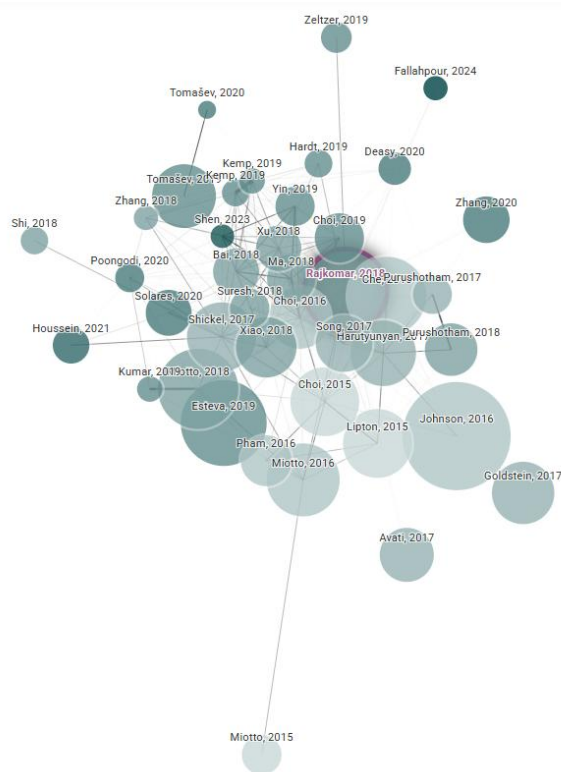
Utilizing timestamps of longitudinal electronic health record data to classify clinical deterioration events	Estados Unidos da América	9	FU et al. 2021
Predicting hypertension onset from longitudinal electronic health records with deep learning	Estados Unidos da América	12	DATTA et al. 2022
Democratizing EHR analyses with FIDDLE: A flexible data-driven preprocessing pipeline for structured clinical data	Estados Unidos da América	82	TANG et al. 2020
Multitask learning and benchmarking with clinical time series data	Estados Unidos da América	628	HARUTYUNYAN et al. 2019
Using machine learning to model older adult inpatient trajectories from electronic health records data	Inglaterra	2	HERRERO-ZAZO et al. 2023
Machine Learning for Decision Support Systems: Prediction of Clinical Deterioration	China	267	SHAMOUT et al. 2023
Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis	Estados Unidos da América	1124	SHICKEL et al. 2018
Data mining information from electronic health records produced high yield and accuracy for current smoking status	Holanda	30	GROENHOF et al. 2020
Machine learning models to detect and predict patient safety events using electronic health records: A systematic review	Iran	15	DEIMAZAR et al. 2023

Fonte: Autor (2025)

Nesta classe de publicações, no **quadro 1**, destaca-se entre outros, o relevante estudo de **RAJKOMAR et al. (2018)**, com 2.227 citações, intitulado *Scalable and accurate deep learning with electronic health records*. Uma pesquisa que derivou diversas outras inspirações como observa-se na **figura 5**, onde os autores sugerem que a modelagem preditiva com dados de prontuários eletrônicos de do paciente (PEP) deverá impulsionar a medicina personalizada e melhorar a qualidade da assistência médica. Propuseram uma representação de todos os prontuários eletrônicos brutos dos pacientes com base no formato *Fast Healthcare Interoperability Resources (FHIR)*. Demonstraram que métodos de aprendizado profundo que utilizam essa representação são capazes de prever com precisão múltiplos eventos médicos de múltiplos centros sem harmonização de dados específicos do local. Validaram uma abordagem usando Registros Eletrônicos de Saúde desidentificados de dois centros médicos acadêmicos dos EUA com 216.221 pacientes adultos hospitalizados por pelo menos 24 horas. Modelos de aprendizado profundo alcançaram alta precisão para tarefas como previsão de: mortalidade hospitalar (área sob a curva do operador do receptor - AUROC - entre os locais 0,93–0,94), readmissão não planejada de 30 dias (AUROC 0,75–0,76), tempo de internação prolongado (AUROC 0,85–0,86) e todos os diagnósticos de alta final de um paciente (AUROC ponderada

por frequência 0,90). Esses modelos superaram os modelos preditivos tradicionais usados clinicamente em todos os casos. Demonstraram que essa abordagem pode ser usada para criar previsões precisas e escaláveis para uma variedade de cenários clínicos. Inteligência artificial: Algoritmo prevê resultados clínicos para pacientes internados em hospitais. Segundo os autores, a inteligência artificial supera os modelos estatísticos tradicionais na previsão de uma série de resultados clínicos de todo o prontuário eletrônico bruto de um paciente. Os autores desenvolveram um *pipeline* de processamento de dados para transformar arquivos de prontuários eletrônicos de saúde (EHR) em um formato padronizado. Em seguida, aplicaram modelos de aprendizado profundo a dados de 216.221 pacientes adultos hospitalizados por pelo menos 24 horas cada em dois centros médicos acadêmicos e demonstraram que seu algoritmo era capaz de prever com precisão o risco de mortalidade, readmissão hospitalar, internação prolongada e diagnóstico de alta. Os autores apresentam um estudo de caso para servir como prova de conceito de como tal algoritmo pode ser usado na prática clínica de rotina no futuro.

**Figura 11 – Trabalhos derivados de: *Scalable and accurate deep learning with electronic health records.***



Fonte: <https://www.connectedpapers.com/> (2025)

**Orientações sobre a legenda e interpretação do gráfico *CONNECT PAPERS*:** Cada círculo, representa um artigo relacionado a publicação original. **A Cor** do círculo: representa o ano de publicação; círculos mais escuros são trabalhos mais recentes; círculos mais claros são trabalhos mais antigos. **Tamanho do círculo:** representa a citação do artigo; círculos maiores são artigos altamente citados e vice-versa. **Clusters** do círculo representa a similaridade dos artigos; quanto mais próximos os *clusters* significam que os artigos têm alta similaridade e vice-versa. **Linha e proximidade:** mostram a semelhança dos papéis; papéis semelhantes estão próximos uns dos outros e com linhas de conexão.

**Quadro 2 – Publicações que abordaram: Deteriorações Clínicas observadas por Aprendizagem de Máquina**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Early prediction of clinical deterioration using data-driven machine-learning modeling of electronic health records	Estados Unidos da América	22	RUIZ et al. 2022
Predicting acute clinical deterioration with interpretable machine learning to support emergency care decision making	Inglaterra	9	BOULITSAKIS LOGOTHETIS et al. 2023
0 of Technologies used by nursing to predict clinical deterioration in hospitalized adults: a scoping review	Brasil	14	GONDIM et al. 2022
Comparison of Multimodal Deep Learning Approaches for Predicting Clinical Deterioration in Ward Patients	Estados Unidos da América	0	KOTULA et al. 2025
Mining clinical phrases from nursing notes to discover risk factors of patient deterioration	Estados Unidos da América	28	KORACH et al. 2020
Systems-Based Physical Assessments: Earlier Detection of Clinical Deterioration and Reduced Mortality	Estados Unidos da América	6	HAMLIN et al. 2023
Importance of specific vital signs in nurses' recognition and response to deteriorating patients: A scoping review	Austrália	9	CONSIDINE et al. 2024
Clinical deterioration: A concept analysis	Estados Unidos da América	76	PADILLA et al. 2018
Predictive Models for In-Hospital Deterioration in Ward Patients	Estados Unidos da América	1	ESCOBAR et al. 2021
Algorithms for Prediction of Clinical Deterioration on the General Wards: A Scoping Review	Holanda	5	PEELEN et al. 2021
Predicting and responding to clinical deterioration in hospitalized patients by using artificial intelligence: Protocol for a mixed method, stepped wedge study	Estados Unidos da América	7	HOLDSWORTH et al. 2021
Nursing Training for Early Clinical Deterioration Risk Assessment: Protocol for an Implementation Study	Brasil	3	LOURENÇO et al. 2023
A scoping review of real-time automated clinical deterioration alerts and evidence of impacts on hospitalised patient outcomes	Austrália	16	BLYTHE et al. 2022

MEWS++: Enhancing the Prediction of Clinical Deterioration in Admitted Patients through a Machine Learning Model	Estados Unidos da América	52	KIA et al. 2020
The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review	Dinamarca	237	BREKKE et al. 2019
Prediction of mortality events of patients with acute heart failure in intensive care unit based on deep neural network	China	7	HUANG et al. 2024
A New Model for tracking and detection of deterioration of vital signs based on artificial neural network	Malasia	2	ABDEL et al. 2019
Use of learning approaches to predict clinical deterioration in patients based on various variables: a review of the literature	Malasia	9	ALSHWAHEEN et al. 2022
Postcardiac Surgery Deterioration Prediction	Estados Unidos da América	17	MATHIS et al. 2022
Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the ICU	Estados Unidos da América	147	KIPNIS et al. 2016

Fonte: Autor (2025)

Nesta classe de publicações, no **quadro 2 – Publicações que abordaram: Deteriorações Clínicas observadas por Aprendizagem de Máquina**, destacamos o relevante estudo de **BREKKE et al. (2019)**, com 237 citações, intitulado *The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review*. Uma pesquisa que derivou diversas outras inspirações como observamos na **Figura 5**, onde os autores apresentaram resultados das 7.366 referências rastreadas, apenas duas foram elegíveis para inclusão. Ambos eram estudos de coorte retrospectivos sem controles. Um examinou a precisão de diferentes modelos de tendência de sinais vitais usando análise de sobrevivência em tempo discreto em 269.999 admissões. Outro incluiu 44.531 admissões médicas examinando a tendência em sinais vitais ponderados pelo *Vitalpac Early Warning Score*. Eles declararam que as tendências dos sinais vitais aumentaram a detecção de deterioração clínica laboratorial. A avaliação crítica foi realizada usando ferramentas de avaliação. Os estudos tinham risco moderado de viés e baixa certeza de evidência. Além disso, quatro estudos examinando tendências em pontuações de alerta precoce, de outra forma elegíveis para inclusão, foram avaliados. Os autores concluíram que existe uma falta de pesquisas em tendências de sinais vitais monitoradas intermitentemente. Os estudos incluídos, embora heterogêneos e imprecisos, indicam um valor agregado da análise de tendências. Isso destaca a necessidade de ensaios bem controlados para avaliar completamente a questão da pesquisa. De modo geral, podemos entender que a análise de deteriorações clínicas por meio de aprendizagem de máquina refere-



**Quadro 3 – Publicações que abordaram: Deteriorações Clínicas observadas por Aprendizagem de Máquina**

Título da Publicação	País de Origem do Estudo	Citações	Autor
Improvement of APACHE II score system for disease severity based on XGBoost algorithm	China	15	LUO et al. 2021
Deep Interpretable Early Warning System for the Detection of Clinical Deterioration	Inglaterra	70	SHAMOUT et al. 2020
Feature Explanations in Recurrent Neural Networks for Predicting Risk of Mortality in Intensive Care Patients	Tailândia	17	NA PATTALUNG et al. 2021

Fonte: Autor (2025)

Nesta classe de publicações, no **quadro 03**, destaca-se o relevante estudo de **SHAMOUT et al. (2020)**, com 70 citações, intitulado *Deep Interpretable Early Warning System for the Detection of Clinical Deterioration*, uma pesquisa que derivou diversas outras inspirações como observamos na **figura 6**, onde os autores demonstraram que as pontuações de alerta precoce são simples de usar, mas consideram os dados como variáveis aleatórias independentes e identicamente distribuídas. Segundo os autores, as aplicações de aprendizado profundo são capazes de aprender com dados sequenciais, porém carecem de interpretabilidade e, portanto, são difíceis de implementar em ambientes clínicos. Propuseram um "Sistema de Alerta Precoce Profundo" chamado de *DEWS*, um modelo de aprendizado profundo de ponta a ponta interpretável que interpola dados temporais e prevê a probabilidade de um evento adverso, definido como o desfecho composto de parada cardíaca, mortalidade ou admissão não planejada na UTI. O modelo foi desenvolvido e validado usando sinais vitais coletados rotineiramente de pacientes internados nos Hospitais da Universidade de Oxford entre 21 de março de 2014 e 31 de março de 2018. Extraídos da base de dados, 45.314 medições de sinais vitais como um conjunto de treinamento balanceado e 359.481 medições de sinais vitais como um conjunto de teste desbalanceado para imitar um cenário real de admissões de emergência. O *DEWS* alcançou precisão superior à do sistema de ponta atualmente implementado em ambientes clínicos, o *National Early Warning Score (NEWS)*, em termos da área total sob a curva característica de operação do receptor (AUROC) (0,880 vs. 0,866) e quando avaliado independentemente para cada um dos três desfechos. Com alta sensibilidade, utilidade clínica aprimorada e maior interpretabilidade, o modelo pode ser facilmente implementado em ambientes clínicos para complementar os sistemas *Early Warning Score (EWS)* existentes. Em



**Quadro 4 – Publicações que abordaram: Inteligência Artificial para Prever Deteriorações Clínicas**

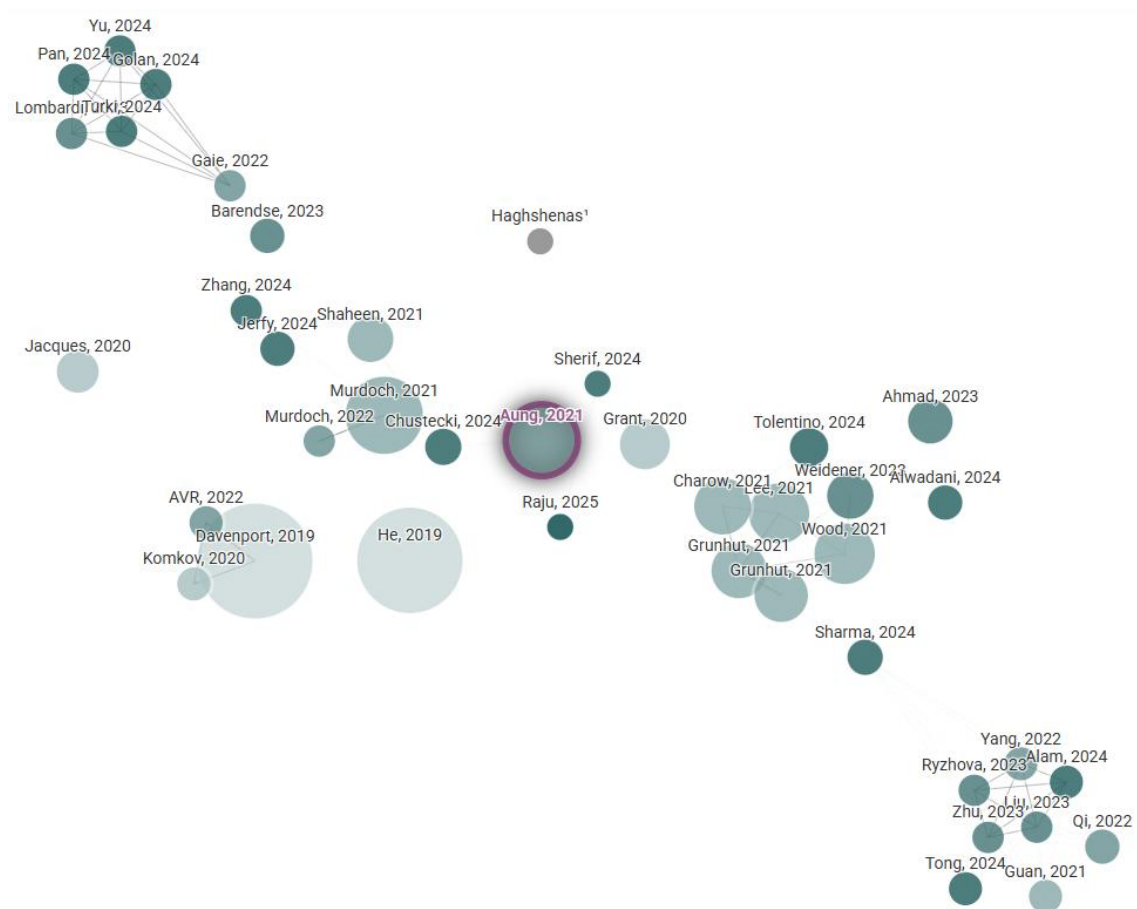
<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Artificial Intelligence for the Prediction of In-Hospital Clinical Deterioration: A Systematic Review	Holanda	19	VELDHUIS et al. 2022
Detecting Patient Deterioration Using Artificial Intelligence in a Rapid Response System	Coreia do Sul	52	CHO et al. 2020
Prediction models using artificial intelligence and longitudinal data from electronic health records: a systematic methodological review	Espanha	25	CARRASCO-RIBELLES et al. 2023
Artificial intelligence in infrastructure construction: A critical review	China	15	CHEN, K et al. 2024
The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare	Inglaterra	246	AUNG et al. 2021
Enhancing In-Hospital Mortality Prediction Using Multi-Representational Learning with LLM-Generated Expert Summaries	Índia	1	BATTULA et al. 2024
Artificial intelligence and clinical deterioration	Australia	8	MALYCHA et al. 2022
Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence	Brasil	51	EVANGELISTA et al. 2020
Integrated multimodal artificial intelligence framework for healthcare applications	Estados Unidos da América	169	SOENKSEN et al. 2022
Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models	Estados Unidos da América	32	CHEN, F. et al. 2024
Systematic review and longitudinal analysis of implementing Artificial Intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain	Australia	17	VEGT et al. 2024
Tecnologias utilizadas pela enfermagem para predição de deterioração clínica em adultos hospitalizados: revisão de escopo	Brasil	2	GONDIM et al. 2022
Foresight—a generative pretrained transformer for modelling of patient timelines using electronic health records: a retrospective modelling study	Inglaterra	43	KRALJEVIC et al. 2024

Fonte: Autor (2025)

Nesta classe de publicações, no **quadro 4**, destacamos o relevante estudo de **AUNG et al. (2021)**, com 246 citações, intitulado *The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare*, uma pesquisa que derivou diversas outras inspirações como observamos na **figura 15**, onde os autores mapearam áreas de convergências e divergências no estudo: **Áreas de acordo:** A IA pode transformar o fluxo de trabalho dos médicos e o atendimento ao paciente por meio de suas aplicações, desde auxiliar

médicos e substituir tarefas administrativas até ampliar o conhecimento médico. **Áreas de divergências:** Desde desafios no treinamento de sistemas de ML até responsabilização pouco clara, a implementação da IA é difícil e, na melhor das hipóteses, gradual. Os médicos também não entendem o que a implementação da IA pode representar. **Pontos de Convergências:** a IA pode, em última análise, ser benéfica na área da saúde, mas exige uma governança metódica semelhante à governança da conduta médica. **Áreas oportunas para o desenvolvimento de pesquisas:** Diretrizes regulatórias são necessárias sobre como implementar e avaliar com segurança a tecnologia de Inteligência Artificial, juntamente com mais pesquisas sobre as capacidades e limitações específicas de seu uso médico. A IA, combinada com dados clínicos, permite uma vigilância contínua e preditiva, ajudando os profissionais a agirem antes que a deterioração clínica se torne grave, melhorando os resultados dos pacientes e a eficiência do cuidado de saúde.

**Figura 14 – Trabalhos derivados de: *The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare***



Fonte: <https://www.connectedpapers.com/> (2025)

### Quadro 5 – Publicações que abordaram: Banco de Dados Utilizados em Saúde

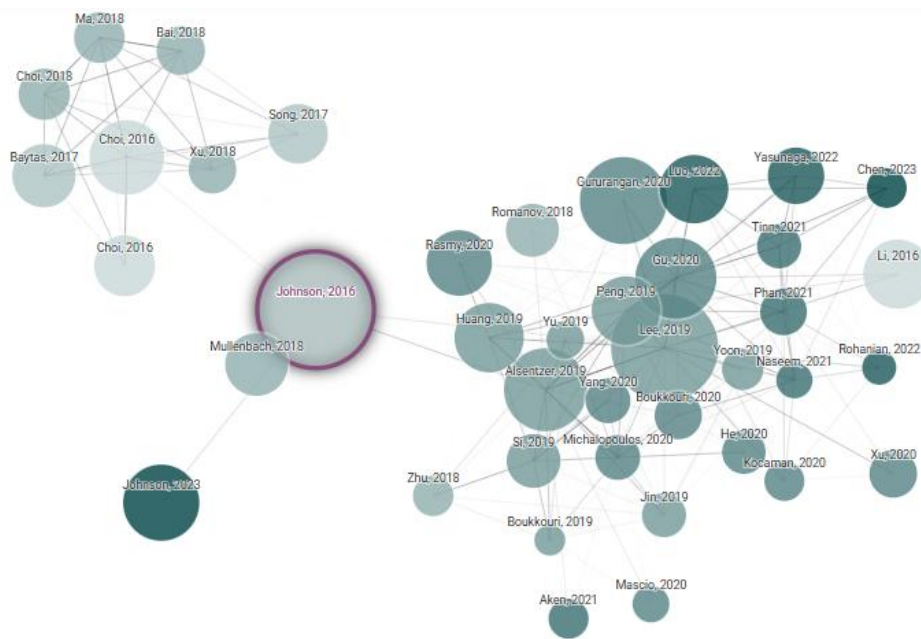
Título da Publicação	Pais de Origem do Estudo	Citações	Autor
Data Quality– and Utility-Compliant Anonymization of Common Data Model–Harmonized Electronic Health Record Data: Protocol for a Scoping Review	Alemanha	1	KAMDJE WABO et al. 2023
MIMIC-IV, a freely accessible electronic health record dataset	Estados Unidos da América	1385	JOHNSON et al. 2023
MIMIC-IF: Interpretability and Fairness Evaluation of Deep Learning Models on MIMIC-IV Dataset	China	30	MENG,et al. 2022
An Extensive Data Processing Pipeline for MIMIC-IV	Índia	44	GUPTA et al. 2022
Interpretability and fairness evaluation of deep learning models on MIMIC-IV dataset	Estados Unidos da América	147	SÃO-JOÃO et al. 2023
Big data and targeted machine learning in action to assist medical decision in the ICU	Estados Unidos da América	48	PIRRACCHIO et al. 2019
Big Data and Data Science in Critical Care	Estados Unidos da América	235	SANCHEZ-PINTO et al. 2018
Big data analytics in adult congenital heart disease: why coding matters	Holanda	6	ASSELBERGS et al. 2019
MIMIC-III, a freely accessible critical care database	Estados Unidos da América	5662	JOHNSON, A et al. 2016
Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations	Italia	2	FANELLI et al. 2023

Fonte: Autor (2025)

Nesta classe de publicações, no **quadro 5 – Publicações que abordaram: Banco de Dados Utilizados em Saúde**, destacamos o exponencial estudo de **JOHNSON, A. et al. (2021)**, com 7263 citações, intitulado *MIMIC-III, a freely accessible critical care database*, uma pesquisa que derivou diversas outras publicações como observamos na **figura 8**, onde os autores fazem reflexões sobre os últimos anos dos banco de dados em saúde, que tem havido um movimento coordenado em direção à adoção de sistemas de registros médicos digitais em hospitais. Nos EUA, por exemplo, o número de hospitais de cuidados intensivos não federais com sistemas digitais básicos aumentou de 9,4% para 75,5% no período de sete anos entre 2008 e 2014. Apesar desse avanço, a interoperabilidade dos sistemas digitais permanece, gerando desafios na integração de dados. Conseqüentemente, o potencial que os dados hospitalares oferecem em termos de compreensão e melhoria do atendimento ainda não foi totalmente explorado. Paralelamente, a comunidade de pesquisa científica vem sendo cada vez mais criticada pela falta de reprodutibilidade dos estudos. Observa-se, nas impressões dos autores, que o lançamento do banco de dados MIMIC-III, uma atualização do amplamente utilizado banco de dados MIMIC-II. O MIMIC-III integra dados clínicos abrangentes e desidentificados

de pacientes internados no Beth Israel *Deaconess Medical Center* em Boston, Massachusetts, e os torna amplamente acessíveis a pesquisadores internacionais por meio de um acordo de uso de dados. A natureza aberta dos dados permite que estudos clínicos sejam reproduzidos e aprimorados de maneiras que, de outra forma, não seriam possíveis.

**Figura 15 – Trabalhos derivados de: *MIMIC-III, a freely accessible critical care database***



Fonte: <https://www.connectedpapers.com/> (2025)

### Quadro 6 – Publicações que abordaram: Entidades Nomeadas com Modelos Pré Treinados

Título da Publicação	Pais de Origem do Estudo	Citações	Autor
Improved characterisation of clinical text through ontology-based vocabulary expansion	Inglaterra	12	SLATER L et al. 2021
Comparative Analysis of Semantic Similarity Techniques for Medical Text	Estados Unidos da América	6	LU H et al. 2022
Systematized Nomenclature of Medicine–Clinical Terminology (SNOMED CT) Clinical Use Cases in the Context of Electronic Health Record Systems: Systematic Literature Review	Finlândia	19	VUOKKO R et al. 2023
Clinical named entity recognition and relation extraction using natural language processing of medical free text: A systematic review	Austrália	32	FRAILE NAVARRO D et al. 2023
Transfer learning for biomedical named entity recognition with neural networks	Canadá	171	GIORGI J et al. 2018

Cross-type biomedical named entity recognition with deep multi-task learning	Estados Unidos da América	252	WANG X et al. 2019
Towards reliable named entity recognition in the biomedical domain	Canadá	74	GIORGI J et al. 2020
BioBART: Pretraining and Evaluation of A Biomedical Generative Language Model	China	132	YUAN H et al. 2020
AMMU: A survey of transformer-based biomedical pretrained language models	Índia	170	KALYAN K et al. 2022
Semantic Similarity Metrics for Evaluating Source Code Summarization	Índia	132	HAQUE S et al. 2022
BioBERT: a pre-trained biomedical language representation model for biomedical text mining	Coreia do Sul	5731	LEE J al. 2020
Modelo de mapeamento semântico entre as terminologias de saúde CID-10 e SNOMED-CT	Brasil	1	GUALDANI F et al. 2024
Obtaining clinical term embeddings from SNOMED CT ontology	Estados Unidos da América	4	ZAHARA F et al. 2024
Automatic full conversion of clinical terms into SNOMED CT concepts	Estados Unidos da América	11	KATE R et al. 2020
Use of the systematized nomenclature of medicine clinical terms (snomed ct) for processing free text in health care: Systematic scoping review	Suíça	47	GAUDET-BLAVIGNAC C et al. 2021
BioBERT Based Named Entity Recognition in Electronic Medical Record	China	54	YU X et al. 2022
Exploring the Value of Pre-trained Language Models for Clinical Named Entity Recognition	Inglaterra	4	BELKADI S et al. 2023
BioBERTpt-A Portuguese Neural Language Model for Clinical Named Entity Recognition	Brasil	75	TERUMI RUBEL et al. 2020
MedBERT: A Pre-trained Language Model for Biomedical Named Entity Recognition	Srilanka	9	VASANTHARAJAN C et al. 2022
BEHRT: Transformer for Electronic Health Records	Inglaterra	469	LI Y et al. 2020

Fonte: Autor (2025)

Na classe de publicações, destacada no **quadro 6**, observa-se o exponencial estudo de **LEE., et al. (2019)** com 5.731 citações, intitulado ***BioBERT: a pre-trained biomedical language representation model for biomedical text mining***, uma pesquisa que derivou diversas outras publicações como observa-se na **figura 9**, onde os autores apresentam o *BioBERT*, um modelo de representação de linguagem pré-treinado para mineração de textos biomédicos. Demonstrado que o pré-treinamento do *BERT* em corpus biomédicos é crucial para sua aplicação no domínio hospitalar. Exigindo modificações arquitetônicas mínimas específicas para cada tarefa. O *BioBERT* supera modelos anteriores em tarefas de mineração de textos biomédicos, como NER, RE e QA. A versão pré-lançada do *BioBERT* em janeiro de 2019, já demonstraram ser muito eficazes em muitas tarefas de mineração de texto biomédico, como as Entidades Nomeadas e notas clínicas de fenótipo-gene humano e clínico temporal. As seguintes versões atualizadas do *BioBERT* estarão disponíveis para a comunidade *bioPNL*: *BioBERT BASE* e *BioBERT LARGE* treinados apenas em resumos do *PubMed* sem



**Quadro 7 – Publicações que abordaram: Instrumentos de Avaliações de Deteriorações Clínicas**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Improvement of APACHE II score system for disease severity based on XGBoost algorithm	China	15	LUO Y et al. 2021
Deep Interpretable Early Warning System for the Detection of Clinical Deterioration	Inglaterra	70	SHAMOUT F et al. 2020
Feature Explanations in Recurrent Neural Networks for Predicting Risk of Mortality in Intensive Care Patients	Tailândia	17	NA PATALLUNG T et al. 2021

Fonte: Autor (2025)

No **quadro 7**, observa-se publicações também conhecidas como: Ferramentas padronizadas utilizadas pelos profissionais de saúde para monitorar, detectar e quantificar manualmente sinais precoces de agravamento do estado clínico dos pacientes. Essas ferramentas geralmente consistem em escadas ou escalas que combinam diferentes parâmetros clínicos, como sinais vitais, estado neurológico, frequência respiratória, saturação de oxigênio e outros indicadores relevantes, atribuindo-lhes pontuações ou categorias que refletem o nível de risco de deterioração. O objetivo desses instrumentos é facilitar a identificação rápida de mudanças adversas no status do paciente, possibilitando intervenções oportunas e prevenindo complicações graves ou óbitos. Entre os exemplos mais comuns estão o MEWS (*Modified Early Warning Score*) e o NEWS (*National Early Warning Score*), que utilizam critérios quantitativos para gerar alertas automáticos ou alertas clínicos, orientando a equipe de saúde na tomada de decisão. Esses instrumentos promovem a padronização na avaliação, aumentam a sensibilidade na detecção precoce de deteriorações e contribuem para uma comunicação mais eficiente entre a equipe multiprofissional, além de orientar protocolos institucionais de resposta rápida. Dessa forma, representam componentes essenciais na prática clínica de monitoração contínua e preventiva em unidades de terapia intensiva, enfermarias e outros ambientes de cuidado hospitalar, contribuindo significativamente para a melhora dos desfechos clínicos dos pacientes.

**Quadro 8 – Publicações que abordaram: Predições de Eventos com Aprendizagem de Máquina**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Predicting the Future — Big Data, Machine Learning, and Clinical Medicine	Estados Unidos da América	66	OBERMEYER Z et al. 2016
Generative AI for Transformative Healthcare: A Comprehensive Study of Emerging Models, Applications, Case Studies, and Limitations	Índia	86	SAI S et al. 2024
Association rules mining for hospital readmission: A case study	Malaysia	8	MISWAN Net al. 2021
A novel deep learning algorithm for real-time prediction of clinical deterioration in the emergency department for a multimodal clinical decision support system	Coreia do Sul	3	CHOI A et al. 2024
A Simulated Prospective Evaluation of a Deep Learning Model for Real-Time Prediction of Clinical Deterioration Among Ward Patients*	Estados Unidos da América	13	SHAH P et al. 2021
A Trend-Based Early Warning Score Can Be Implemented in a Hospital Electronic Medical Record to Effectively Predict Inpatient Deterioration	Australia	15	BELL D et al. 2021
Predictive modeling in urgent care: A comparative study of machine learning approaches	Estados Unidos da América	43	TANG F et al. 2018
A Meta-embedding-based Ensemble Approach for ICD Coding Prediction	India	2	RAJENDRAN P et al. 2018
Predicting cardiac arrest and respiratory failure using feasible artificial intelligence with simple trajectories of patient data	Coreia do Sul	50	KIN J et al. 2019
Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost	China	387	HOU N et al. 2020
Combining structured and unstructured data for predictive models: a deep learning approach	Estados Unidos da América	191	ZHANG D et al. 2020
Predicting Intensive Care Unit Length of Stay and Mortality Using Patient Vital Signs: Machine Learning Model Development and Validation	Estados Unidos da América	60	ALGHATANI K et al. 2020
Intensive Care Unit Mortality Prediction: An Improved Patient-Specific Stacking Ensemble Model	Egito	82	EL-RAASHIDY N et al. 2020
Integrating Structured and Unstructured Data for Predicting Emergency Severity: A Combined Predictive and Association Study Using PNL Models	China	1	ZHANG XChu et al. 2024
Predicting Patient Deterioration: A Review of Tools in the Digital Hospital Setting	Estados Unidos da América	38	MANN K et al. 2021
eDoctor: machine learning and the future of medicine	Inglaterra	596	HANDELMAN G et al. 2018
An early sepsis prediction model utilizing machine learning and unbalanced data processing in a clinical context	China	2	Edelson et al. 2021
A Long Short-Term Memory Ensemble Approach for Improving the Outcome Prediction in Intensive Care Unit	China	34	XIA J et al. 2019

Using deep learning with attention mechanism for identification of novel temporal data patterns for prediction of ICU mortality	Estados Unidos da América	6	GE W et al. 2022
DeepSigns: A predictive model based on Deep Learning for the early detection of patient health deterioration	Brasil	40	DA SILVA D et al. 2021
Using electronic health record collected clinical variables to predict medical intensive care unit mortality	Estados Unidos da América	69	CALVERT J et al. 2016
Multilayer dynamic ensemble model for intensive care unit mortality prediction of neonate patients	Coreia do Sul	29	JURAEV F et al. 2022
Automated data mining of the electronic health record for investigation of healthcare-associated outbreaks	Estados Unidos da América	12	SUNERMANN A et al. 2019

Fonte: Autor (2025)

No **quadro 8**, temos publicações conhecidas como algoritmos de inteligência artificial para antecipar a ocorrência de acontecimentos clínicos futuros com base na análise de dados históricos e em tempo real de pacientes. Esses sistemas utilizam modelos preditivos treinados em grandes conjuntos de dados, incluindo sinais vitais, exames laboratoriais, registros eletrônicos de saúde, informações demográficas e outras variáveis clínicas relevantes, para identificar padrões e tendências que antecedem eventos específicos, como deterioração, sepse, insuficiência respiratória, falência de órgãos, complicações cirúrgicas ou desfechos adversos. A aprendizagem de máquina emprega algoritmos como redes neurais profundas, árvores de decisão, support vector machines (SVM) e modelos ensemble, capazes de capturar relações complexas e não lineares entre variáveis, fornecendo probabilidades ou scores que indicam o risco de esses eventos acontecerem em um horizonte de tempo definido. Essas previsões permitem que equipes de saúde realizem intervenções preventivas, adotem estratégias de monitoramento intensificado ou iniciem tratamentos precoces, aumentando a eficácia do cuidado clínico e melhorando os desfechos dos pacientes. A implementação dessas tecnologias requer validação rigorosa, integração com sistemas de informação hospitalar e capacitação da equipe, de modo a garantir a confiabilidade, interpretabilidade e utilidade clínica das previsões geradas. Assim, as previsões de eventos com aprendizagem de máquina representam uma ferramenta avançada na medicina preditiva, contribuindo para o cuidado personalizado e a gestão proativa de riscos clínicos.

**Quadro 9 – Publicações que abordaram: Redes Neurais usadas em Registros Eletrônicos de Saúde**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Preparing for Success in Healthcare Information and Management Systems	Estados Unidos da América	7	HIMSS et al. 2021
LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy	Estados Unidos da América	1	KOWSHER M et al. 2021
Critical Data for Critical Care: A Primer on Leveraging Electronic Health Record Data for Research From Society of Critical Care Medicine's Panel on Data Sharing and Harmonization	Estados Unidos da América	1	HEAVNER S et al. 2024
Pneumonia Outcome Prediction Using Structured And Unstructured Data From EHR	Japão	10	MUGISHA C et al. 2020
Using a Multi-Task Recurrent Neural Network With Attention Mechanisms to Predict Hospital Mortality of Patients	China	45	YU R et al. 2022
Continuous and automatic mortality risk prediction using vital signs in the intensive care unit: a hybrid neural network approach	Australia	34	BAKER S et al. 2020
Deep learning for electronic health records: A comparative review of multiple deep neural architectures	Inglaterra	179	AYALA SOLARES J et al. 2024
Piloting electronic medical record-based early detection of inpatient deterioration in community hospitals	Estados Unidos da América	92	ESCOBAR G et al. 2016
Análise do uso de ferramentas de registro em prontuário eletrônico em Unidades de Terapia Intensiva	Brasil	1	COUTINHO J et al. 2022
Dynamic and explainable machine learning prediction of mortality in patients in the intensive care unit: a retrospective study of high-frequency data in electronic patient records	Dinamarca	209	THORSEN-MEYER H et al. 2020
A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures	China	3818	YU Y et al. 2019
Learning to Diagnose with LSTM Recurrent Neural Networks	Estados Unidos da América	1112	LIPTON Z et al. 2020
Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long shortterm memory	Estados Unidos da América	152	LIN Y et al. 2021
Evolving Deep Neural Networks	Estados Unidos da América	892	MIKKULAINEN R et al. 2020
A Novel and Reliable Framework of Patient Deterioration Prediction in Intensive Care Unit Based on Long Short-Term Memory-Recurrent Neural Network	Malasia	14	ALSHWAHEEN T et al. 2021

Fonte: Autor (2025)

Os temas, abordados no **quadro 9**, são conhecidos como modelos de aprendizagem de máquina inspirados na estrutura do sistema nervoso biológico, capazes de aprender e generalizar padrões complexos presentes nos dados clínicos armazenados em sistemas de informação hospitalares. Essas redes consistem em camadas de nós (neurônios artificiais) conectados por enlaces ponderados que processam informações provenientes de registros eletrônicos de saúde, incluindo sinais vitais, exames laboratoriais, anotações clínicas, históricos médicos e outros dados relacionados ao paciente. Ao serem treinadas com grandes volumes de dados históricos, as redes neurais ajustam seus pesos internos para reconhecer associações e tendências relevantes, possibilitando tarefas como classificação de diagnósticos, previsão de eventos adversos, detecção de anomalias, segmentação de recursos e avaliação da evolução clínica do paciente. Sua capacidade de modelar relações não lineares e de integrar múltiplas fontes de dados torna as redes neurais uma ferramenta poderosa na análise de registros eletrônicos, promovendo uma abordagem mais precisa, automatizada e personalizada na tomada de decisões clínicas. A aplicação dessas redes exige procedimentos de validação rigorosos, para evitar *overfitting* e assegurar sua interpretabilidade e confiabilidade em ambientes de alta complexidade, além de uma adequada integração com sistemas de suporte à decisão clínica e interfaces de usuário amigáveis. Dessa forma, as redes neurais representam uma tecnologia emergente e promissora na gestão e análise de dados de saúde eletrônicos, contribuindo para otimizar o cuidado, melhorar os desfechos pacientes e facilitar a implementação de cuidados baseados em evidências.

**Quadro 10 – Publicações que abordaram: Processamento de Linguagem Natural utilizadas em Registros Eletrônicos de Saúde**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Biomedical Text Normalization through Generative Modeling	Estados Unidos da América	0	BERKOWITZ J et al. 2024
Cohort identification From Free-Text Clinical Notes Using Snomed CTS Semantic Relations	Estados Unidos da América	3	CHANG E et al. 2022
An annotated association mining approach for extracting and visualizing interesting clinical events	Estados Unidos da América	15	SHRESTHA A et al. 2021
A survey on recent named entity recognition and relationship extraction techniques on clinical texts	Estados Unidos da América	125	BOSE P et al. 2021
Transformers-sklearn: a toolkit for medical language understanding with transformer-based models	China	32	YANG F et al. 2021
A Review on Electronic Health Record Text-Mining for Biomedical Name Entity Recognition in Healthcare Domain	China	32	PIR NOMAN AHMAD et al. 2023

Biomedical and clinical English model packages for the Stanza Python PNL library	Estados Unidos da América	128	ZHANG Y et al. 2021
Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications	Estados Unidos da América	1924	SAVOVA G et al. 2010
Use of unstructured text in prognostic clinical prediction models: a systematic review	Holanda	61	SEINEN T et al. 2018
Towards unstructured mortality prediction with free-text clinical notes	Canadá	39	HASHIR M et al. 2022
ISeeU2: Visually interpretable mortality prediction inside the ICU using deep learning and free-text medical notes	Nova Zelândia	17	CAICEDO-TORRES et al. 2020
Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances	Inglaterra	189	VELUPILLAI S al. 2023
Integrating structured and unstructured data for predicting emergency severity: an association and predictive study using transformer-based natural language processing models	Estados Unidos da América	26	CHIU C et al. 2024
Neural Natural Language Processing for unstructured data in electronic health records: A review	Estados Unidos da América	160	LI I et al. 2019
Identifying risks areas related to medication administrations - text mining analysis using free-text descriptions of incident reports	Finlândia	22	HARKANEN M et al. 2021
Text-mining in electronic healthcare records can be used as efficient tool for screening and data collection in cardiovascular trials: a multicenter validation study	Holanda	30	VAN DIJK W et al. 2023
An analysis of Chinese nursing electronic medical records to predict violence in psychiatric inpatients using text mining and machine learning techniques	Taiwan	11	HU Y et al. 2024
Data processing and text mining technologies on electronic medical records: A review	China	155	SUN W et al. 2022
Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review	Estados Unidos da América	301	KOLECK T et al. 2022
Using a Text Mining Approach to Explore the Recording Quality of a Nursing Record System	Taiwan	9	SHANG H et al. 2018

Fonte: Autor (2025)

Os autores citados no **quadro 10**, sugerem que o processamento de linguagem natural (PLN) utilizado em registros eletrônicos de saúde refere-se à aplicação de técnicas de inteligência artificial e linguística computacional para analisar, interpretar, extrair e organizar informações contidas em textos clínicos não estruturados, como anotações médicas, relatos de pacientes, laudos de exames e outras documentações subjetivas. Essas técnicas envolvem etapas como segmentação de texto, reconhecimento de entidades clínicas (como medicamentos, diagnósticos, procedimentos), normalização de termos médicos, análise sintática e semântica, além de extração automática de informações relevantes e modulação de linguagem natural para

gerar resumos, alertas ou relatórios estruturados. O objetivo do PLN na área da saúde é transformar dados textuais desestruturados em informações quantitativas e qualitativas acessíveis a sistemas de apoio à decisão clínica, melhorar a interoperabilidade entre sistemas, facilitar revisões de registros, identificar padrões de risco e auxiliar na monitoração de condições de saúde. Para isso, utiliza algoritmos de aprendizado de máquina, modelos de linguagem, ontologias médicas e dicionários especializados que garantem maior precisão na interpretação de termos clínicos específicos e jargões médico-legais.

### **Quadro 11 – Publicações que abordaram: Outras Classes relacionadas ao Aprendizagem de Máquina**

<b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Biomedical named entity recognition and linking datasets: Survey and our recent development	Taiwan	40	HUANG M et al. 2020
Supporting rapid innovation in research data capture and management: the REDCap external module framework	Estados Unidos da América	2	SHENG A et al. 2025
ChatGPT for Education and Research: Opportunities, Threats, and Strategies	Bangladesh	617	RAHMAN M et al. 2018

Fonte: Autor (2025)

Outros exemplos de artigos descritas no **quadro 11**, nos direciona entender também que o uso da aprendizagem de máquina (AM) aplicada em registros eletrônicos de saúde (RES) também refere-se ao uso de algoritmos de inteligência artificial que são treinados com grandes volumes de dados clínicos digitais para identificar padrões, realizar classificações preditivas, automatizar processos de análise e gerar insights clínicos de maneira eficiente e precisa. Esses algoritmos, incluindo redes neurais profundas, máquinas de vetores de suporte, árvores de decisão e modelos ensemble, permitem a extração de informações relevantes a partir de conjuntos de dados heterogêneos, como sinais vitais contínuos, laudos, anotações de texto livre, imagens e dados genômicos, além de auxiliar na predição de eventos adversos, na classificação de diagnósticos e na personalização do tratamento. Nos próximos 10 anos, as tendências na aplicação da aprendizagem de máquina em RES apontam para uma integração maior de modelos multimodais, capazes de combinar diferentes tipos de dados clínicos, a adoção de sistemas de aprendizado contínuo (online learning) que se atualizam em tempo real com as novas informações, e o desenvolvimento de algoritmos interpretáveis (explicáveis) que promovam maior confiança por parte dos profissionais de saúde. Além disso, espera-se uma

maior incorporação de técnicas de aprendizagem federada para preservar a privacidade dos dados, a implementação de modelos preditivos proativos que antecipem deteriorações clínicas com elevada acurácia, e a utilização de inteligência artificial explicável para melhorar a aceitação clínica e facilitar a tomada de decisão informada. Essas tendências impulsionarão a automação, precisão e personalização do cuidado de saúde, promovendo uma medicina cada vez mais preditiva, preventiva e orientada por dados, contribuindo significativamente para melhorias nos desfechos clínicos e na gestão de recursos hospitalares.

**Quadro 12 – Publicações que abordaram: Revisões Sistemáticas relacionadas ao  
Aprendizagem de Máquina**

Guiding Principles for Data <b>Título da Publicação</b>	<b>Pais de Origem do Estudo</b>	<b>Citações</b>	<b>Autor</b>
Results of a Systematic Review and Modified Delphi From the Society of Critical Care Medicine Data Science Campaign	Estados Unidos da América	1	CHEN et al. 2025
Changing patterns in reporting and sharing of review data in systematic reviews with meta-analysis of the effects of interventions: Cross sectional meta-research study	Australia	31	NGUYEN P et al. 2023
Guidance to best tools and practices for systematic reviews	Estados Unidos da América	79	KOLASKI K et al. 2024

Fonte: Autor (2025)

A síntese dos achados das publicações citadas no **quadro 12**, demonstram que o aprendizado de máquina (AM) pode ser também empregado para otimizar as revisões sistemáticas com meta-análises de registros eletrônicos de saúde (RES) ao automatizar e aprimorar várias etapas do processo de revisão, incluindo busca, seleção, extração de dados, avaliação de qualidade e síntese de evidências. Algoritmos de processamento de linguagem natural (PLN) podem ser utilizados para realizar buscas automatizadas e realizar triagens preliminares de grandes volumes de literatura clínica e dados de RES, identificando de forma rápida e precisa estudos relevantes e extrair informações estruturadas de textos livres e laudos médicos, reduzindo significativamente o tempo e o esforço manual. Modelos de aprendizado supervisionado podem classificar a relevância e a qualidade dos estudos científicos, enquanto técnicas de aprendizado não supervisionado, como *clustering*, podem identificar padrões e subgrupos nos dados de RES, facilitando a análise de grande escala. Além disso, métodos de aprendizado de máquina podem ser utilizados para padronizar e automatizar a extração de

variáveis clínicas, ajustar critérios de inclusão e exclusão, e ajudar na avaliação de risco de viés, aumentando a robustez das revisões sistemáticas. Nos próximos 10 anos, as tendências indicam uma maior integração de sistemas de inteligência artificial com plataformas de revisão, possibilitando uma atualização contínua de evidências em tempo real, uso de algoritmos explicáveis para garantir transparência nas conclusões, e o desenvolvimento de *pipelines* automatizados de meta-análise, que combinem dados de RES com outros conjuntos de dados clínicos e genômicos, possibilitando análises mais abrangentes e personalizadas. Além disso, espera-se uma maior adoção de técnicas de mineração de dados avançadas, aprendizado federado para proteger a privacidade, e frameworks integrados de inteligência artificial que possam evoluir de forma autônoma, fornecendo suporte em decisões clínicas, identificando lacunas de evidência, e promovendo uma medicina baseada em dados mais rápida, eficiente e confiável.

### Quadro 13 – Publicações que abordaram: Estudos com Método Prisma

Título da Publicação	Pais de Origem do Estudo	Citações	Autor
PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews	Australia	7239	PAGE M et al. 2022
PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis	Africa do Sul	543	HADDAWAY N et al. 2022
Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement	Canadá	14599	MOHER D et al. 2018

Fonte: Autor (2025)

No **quadro 13**, são observadas reflexões sobre o uso de diretrizes como o PRISMA (*Preferred Reporting Items for Scoping Reviews*) e representam elementos essenciais na condução de revisões sistemáticas de alta qualidade, especialmente no contexto de análises de registros eletrônicos de saúde (RES). Essas abordagens visam garantir transparência, reprodutibilidade e abrangência na síntese de evidências, permitindo uma compreensão estruturada das atuais lacunas, tendências e desafios na utilização de RES para diferentes propósitos clínicos e de pesquisa. O PRISMA oferece um framework padronizado para relatar processos de busca, seleção, avaliação e análise de estudos, minimizando viés e facilitando a

avaliação crítica por parte da comunidade científica. As revisões de escopo, por sua vez, expandem a análise para mapear um amplo espectro de evidências, identificar áreas emergentes e estratégias inovadoras de uso de RES, além de orientar futuras direções de pesquisa e implementação tecnológica.

Em futuro próximo, espera-se uma incorporação mais robusta de ferramentas de automação e inteligência artificial na condução dessas revisões, envolvendo desde a triagem automatizada de estudos e extração de dados até o mapeamento de tendências por meio de técnicas de análise de grande volume de textos, como PLN e análise de redes. Todavia, diversos desafios persistirão, incluindo questões relacionadas à heterogeneidade dos dados de RES, problemas de interoperabilidade, privacidade e ética na utilização de dados sensíveis, além da necessidade de maior padronização dos formatos e critérios de validação.

### 3.0 MATERIAIS E MÉTODOS

#### 3.1 CARACTERIZAÇÃO DA NATUREZA DE PESQUISA

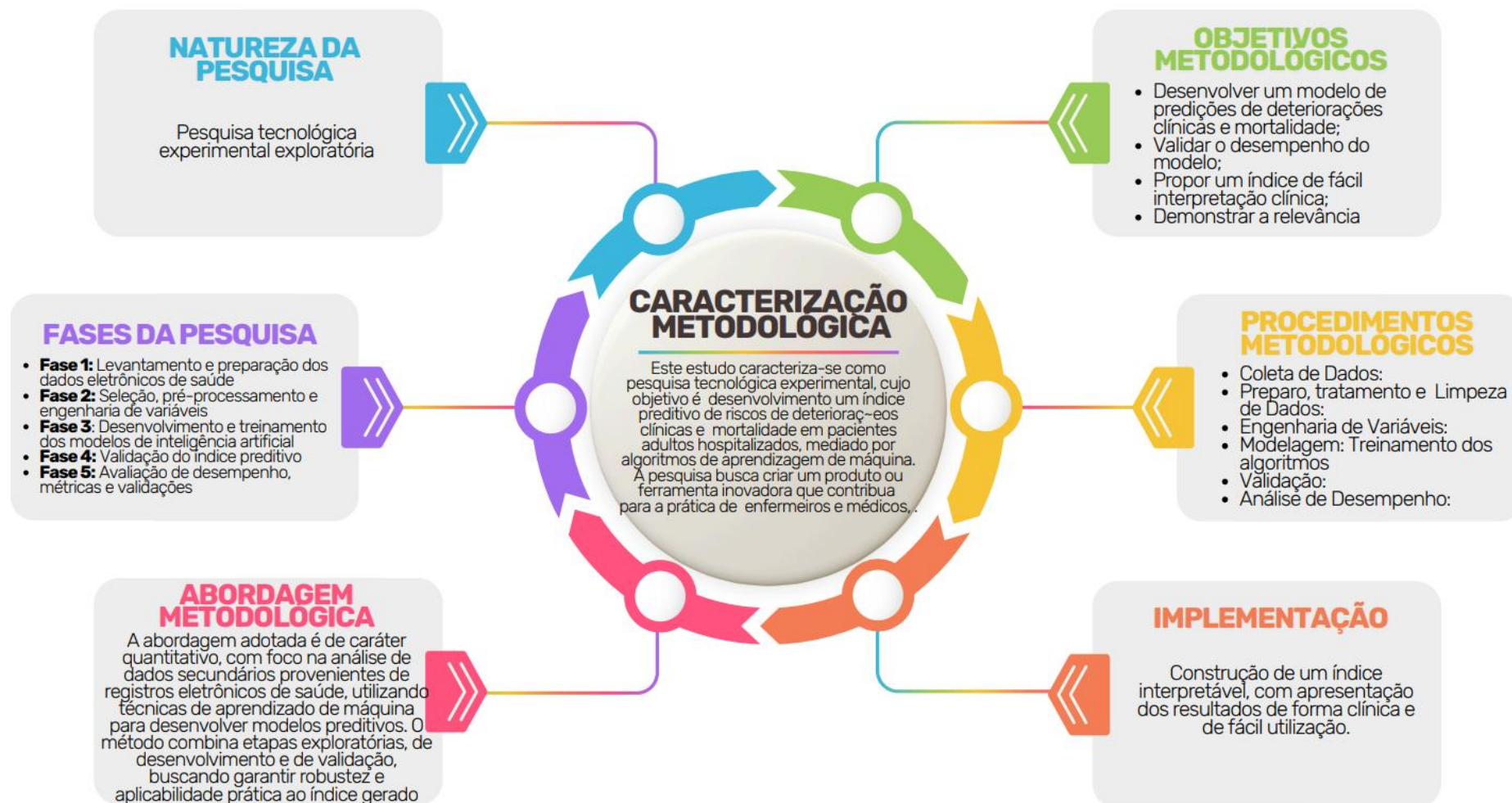
Nesta seção serão descritos os materiais e métodos usados para realização da pesquisa. O estudo é de natureza tecnológica experimental exploratória;

Na **figura 17**, observa-se uma ilustração que faz referência a natureza da pesquisa científica, caracteriza-se por seu caráter sistemático, racional e investigativo, com o objetivo de gerar conhecimento confiável e válido que contribua para o avanço do entendimento em uma determinada área do saber. Este processo é sequencial e organizado, dividido em fases bem distintas que envolvem desde a formulação do problema até a divulgação dos resultados, passando pela coleta, análise e interpretação dos dados. Essas fases garantem a integridade metodológica e a coerência do estudo, permitindo uma construção sólida de conhecimento fundamentado em evidências empíricas.

A abordagem metodológica adotada na pesquisa refere-se ao conjunto de estratégias e técnicas que orientam a investigação e deve ser selecionada de acordo com a natureza do problema, os objetivos almejados e o tipo de dado que se pretende obter. Pode-se assumir diferentes formas, como a qualitativa, que busca compreender fenômenos complexos e subjetivos, ou a quantitativa, que usa dados numéricos para análise estatística, ou ainda uma abordagem mista, que combina elementos de ambas. A escolha adequada da abordagem é fundamental para garantir a consistência do método e a relevância dos resultados.

Os objetivos metodológicos orientam o desenvolvimento da pesquisa, podendo ser classificados em exploratórios, que visam descobrir novas informações; descritivos, que detalham as características de um fenômeno; ou explicativos, que buscam estabelecer relações de causa e efeito. Esses objetivos estabelecem o roteiro que guiará as ações futuras, definindo claramente o que se pretende alcançar e quais hipóteses ou questões serão respondidas ao longo do estudo. Assim, eles influenciam diretamente na escolha dos procedimentos metodológicos e na estrutura do desenvolvimento do trabalho.

Figura 17 – Caracterização Metodológica da Pesquisa



Fonte: Autor (2025)

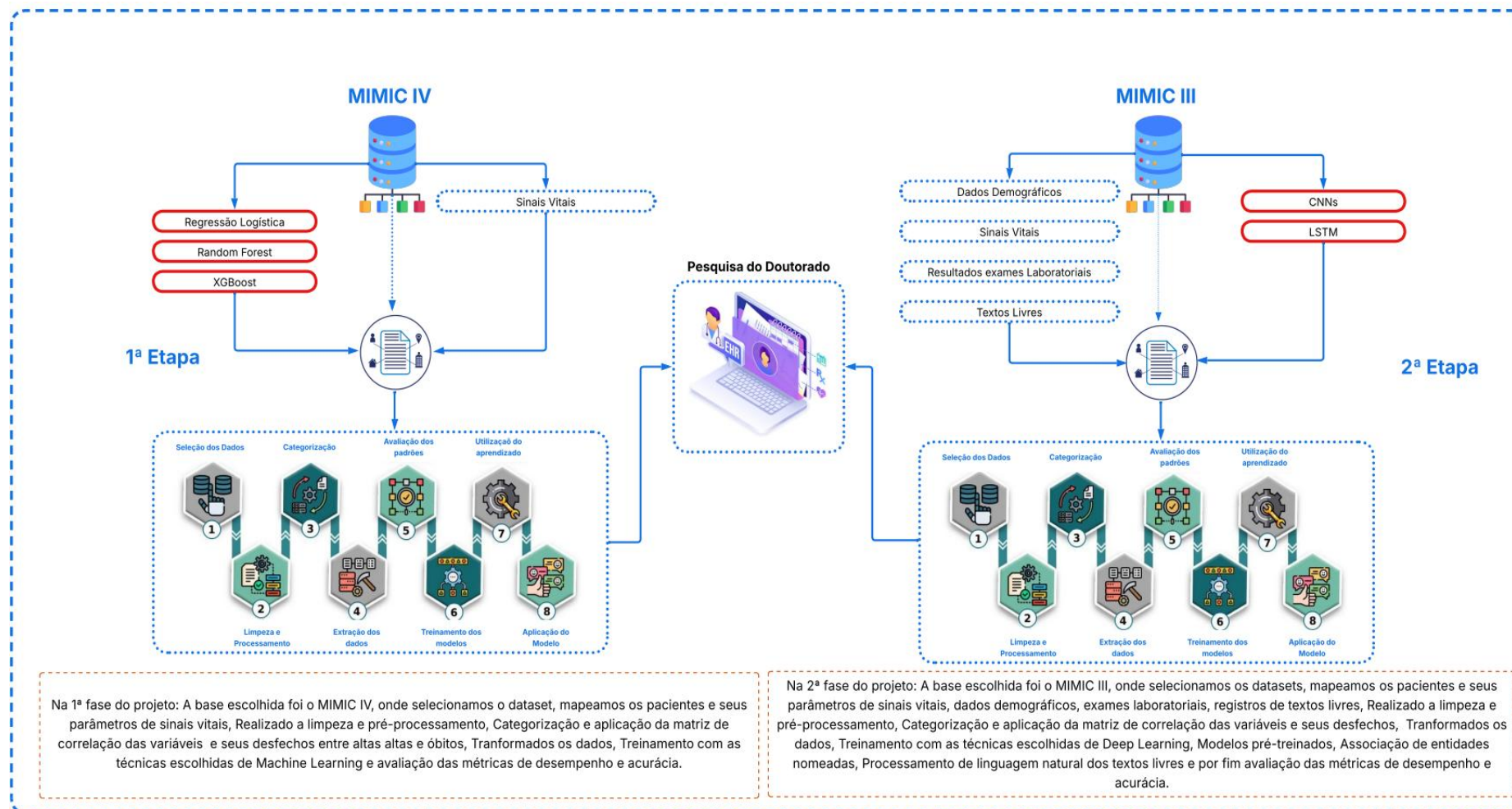
Os procedimentos metodológicos envolvem o conjunto de etapas e técnicas utilizadas para coletar, processar e analisar os dados, destacando-se a definição dos instrumentos de coleta, a amostragem, os critérios para validação e confiabilidade, além das técnicas de análise que garantam a integridade das conclusões. A precisão na execução desses procedimentos é essencial para assegurar a credibilidade da pesquisa e a reprodutibilidade dos resultados, bem como para reforçar a validade das inferências feitas a partir dos dados obtidos.

Desta forma, na etapa de implementação de uma pesquisa aplicada, o foco recai na aplicação prática do conhecimento produzido, com o objetivo de solucionar problemas específicos ou promover melhorias em contextos reais. Essa etapa demanda planejamento detalhado, acompanhamento sistemático e ajustes contínuos, visando à consecução de resultados concretos que possam gerar impacto social, econômico ou tecnológico. A pesquisa aplicada representa a ponte entre a teoria e a prática, colocando o saber científico em ação para promover transformações efetivas no ambiente de atuação, além de validar a utilidade prática dos conhecimentos gerados e contribuir para o desenvolvimento de soluções inovadoras e sustentáveis.

### **3.2 DESCRIÇÃO DAS PRINCIPAIS ETAPAS DA PESQUISA**

Esta pesquisa é parte integrante de uma sequência de experimentos computacionais que produzidos durante a elaboração de uma tese de doutorado pelo Programa de Mestrado e Doutorado em Informática e Gestão do conhecimento na Universidade Nove de Julho de São Paulo, Brasil. A condução do presente estudo foi organizada em duas macros etapas, conforme ilustradas na **figura 18**.

**Figura 18: Exibição do planejamento das principais macro etapas usada na metodologia da pesquisa**



Fonte: Autor (2025)

Tais fases, e as suas respectivas etapas, são mais bem descritas a seguir:

- **Fase 1:** procurou-se demonstrar no início da pesquisa, a partir de informações de registros eletrônicos de saúde da base de dados MIMIC-IV, que apenas os sinais vitais dos pacientes, em combinação com técnicas de aprendizagem de máquina do tipo Regressão Logística, Random Forest e XGBoost, eram capazes de classificar o desfecho de um paciente adulto hospitalizado (óbito ou alta). Os resultados foram publicados em um primeiro artigo científico.
- **Fase 2:** nesta segunda etapa foi realizada uma revisão de escopo, uma análise dos atributos, bem como o cálculo do índice de deterioração para predição futura da condição do paciente, a criação do banco de dados com os indicadores para milhares de pacientes e os experimentos com os algoritmos de aprendizagem de máquina, bem como o teste, após o treinamento do modelo com dados da base MIMIC-III, com dados da base MIMIC-IV, tal fase pode ser dividida em 4 etapas a saber:

- **Etapa 2.1:** nesta etapa foi produzida uma revisão de escopo, seguindo as diretrizes de PRISMA (*-Preferred Reporting Items for Systematic reviews and Meta-Analyses*), como sugere (MOHER et al., 2016, KOLASKI et al., 2023, NGUYEN et al., 2022, HADDAWAY et al., 2022 e PAGE et al., 2021);

- **Etapa 2.2:** nesta etapa do trabalho foram analisadas estatisticamente as variáveis clínicas e de laboratório mais relevantes extraídas do banco de dados MIMIC-III e determinado o cálculo do índice para a predição clínica futura em pacientes adultos com base nas grandezas clínicas e laboratoriais identificadas como mais importantes para o futuro desfecho;

- **Etapa 2.3:** na terceira etapa estudo, foram construídos o banco de dados com os índices de milhares de pacientes e os algoritmos de aprendizagem de máquina para a predição temporal futura do índice de deterioração dos pacientes.

- **Etapa 2.4:** na quarta etapa do estudo foi testado o melhor modelo de aprendizagem de máquina treinado, usando os dados do MIMIC-III, com dados do índice a partir de pacientes da base MIMIC-VI.

A seguir será realizado o detalhamento de cada etapa do estudo referente à Fase 2, ou seja, das Etapas 2.1, 2.2, 2.3 e 2.4.

### 3.2.1 DETALHAMENTO DA ETAPA 2.1 DA FASE 2

Na etapa **2.1**, para a revisão de escopo, de acordo com as propostas de **KAMDIE et al. (2023)**, foram utilizadas e consultadas 6 das principais bases de dados, que indexam produções relacionadas a temática da pesquisa, como: *Medical Literature Analysis and Retrieval System online* (MEDLINE), **Excerpta Medica DataBASE** (EMBASE), *The Cumulative Index to Nursing and Allied Health Literature* (CINAHL) via plataforma EBSCO, *Scopus* (Elsevier) e *Web of Science* e **IEEE Xplore**.

As bases foram consultadas no período de 01 de janeiro de 2025 a 30 de Julho de 2025 para publicações em inglês e português de títulos resumos com palavras chaves como: (1) Índice de Deteriorações clínicas, (2) Mortalidade, (3) Registros Eletrônicos de Saúde, (4) Aprendizado de Máquina, (5) Séries Temporais, (6) Inteligência Artificial. Or *abstract titles with keywords such as: (1) Clinical Deteriorations Index, (2) Mortality, (3) Electronic Health Records, (4) Machine Learning, (5) Times Series, (6) Artificial Intelligence.*

Cabe destacar ainda que, dado o objetivo principal da pesquisa, o trabalho empregou uma estratégia voltada para identificar artigos e *papers* com foco em pesquisas de implementações de algoritmos de aprendizado de máquina (PIAAM) em Registros Eletrônicos de Saúde (RES).

### 3.2.2 ESTRATÉGICAS DE PESQUISAS NA FASE DE FUNDAMENTAÇÃO

Esta pesquisa é parte integrante de uma sequência de experimentos computacionais que produzidos durante a elaboração de uma tese de doutorado pelo Programa de Mestrado e Doutorado em Informática e Gestão do conhecimento na Universidade Nove de Julho de São Paulo, Brasil. A condução do presente estudo, foi organizada em duas etapas:

A **Etapa 1**: Nesta fase foi produzida uma revisão de escopo, seguindo as diretrizes de PRISMA - *Preferred Reporting Items for Systematic reviews and Meta-Analyses*), como sugere (**MOHER et al., 2016, KOLASKI et al., 2023, NGUYEN et al., 2022, HADDAWAY et al., 2022, KAMDIE et al. (2023) e PAGE et al., 2021**).

A **Etapa 2**: na segunda etapa do estudo, foram construídos os experimentos com elaboração dos algoritmos de Aprendizado de Máquina.

Foram utilizadas e consultadas 6 das bases de dados, que indexam produções

relacionadas a temática da pesquisa, como: *Medical Literature Analysis and Retrieval System online* (MEDLINE), *Excerpta Medica DataBASE* (EMBASE), *The Cumulative Index to Nursing and Allied Health Literature* (CINAHL) via plataforma EBSCO, *Scopus* (Elsevier) e *Web of Science* e *IEEE Xplore*.

**Tabela 2: Tabela com a descrição das Palavras Chaves/Mesh**

Palavras Chaves / Mesh		
Ordem	Termos em Inglês	Termos em Português
1	<i>Clinical Deteriorations Index</i>	Índice de Deteriorações Clínicas
2	<i>Mortality</i>	Mortalidade
3	<i>Electronic Health Records</i>	Registros Eletrônicos de Saúde
4	<i>Machine Learning</i>	Aprendizagem de Máquina
6	<i>Times Series</i>	Séries Temporais
8	<i>Artificial Intelligence</i>	Inteligência Artificial

Fonte: Autor (2025)

### 3.2.3 DETALHAMENTO DA ETAPA 2.2 DA FASE 2

Ainda na **2ª macro fase**, mais particularmente na etapa 2.2, iniciou-se o estudo das variáveis mais relevantes contidas na base de dados utilizada, qual seja, a MIMIC-III, por sua maior robustez de corpus e ampla oferta de atributos, onde são utilizados dos registros eletrônicos de saúde, além dos sinais vitais, dados demográficos, resultados de exames laboratoriais e textos livres produzidos por Enfermeiros e Médicos. Tal base de dados, também utilizada por **HOU et al. (2020)**, está disponível publicamente, e é originado dos registros eletrônicos de saúde do Beth Israel Deaconess Medical Center (BIDMC), gerenciados pelo MIT (*Massachusetts Institute of Technology*). (**JOHNSON et al., 2016**)

### DETALHAMENTO TÉCNICO DO DATASET MIMIC-III

#### 1. Escopo e Origem dos Dados

- **Instituição:** Beth Israel Deaconess Medical Center (Boston, EUA).
- **Período:** Coleta realizada entre 2001 e 2012.
- **Sistemas de Origem:** Os dados provêm de dois sistemas de prontuário eletrônico

distintos: CareVue (2001-2008) e MetaVision (2008-2012).

## 2. Estatísticas do Banco de Dados (v1.4)

- **Pacientes únicos:** Aproximadamente 46.520 indivíduos.
- **Admissões hospitalares:** Cerca de 58.976 registros.
- **População:** Inclui adultos (38.597) e recém-nascidos (7.870).

## 3. Estrutura e Tabelas

O banco é composto por **26 tabelas** em formato CSV, interconectadas por identificadores únicos (SUBJECT\_ID, HADM\_ID, ICUSTAY\_ID). As tabelas são divididas em categorias:

- **Dados de Pacientes:** PATIENTS (demografia), ADMISSIONS (internações) e ICUSTAYS.
- **Eventos Clínicos:** CHARTEVENTS (sinais vitais, observações), LABEVENTS (exames laboratoriais) e OUTPUTEVENTS (balanço hídrico).
- **Dicionários:** Tabelas prefixadas com D\_ (ex: D\_ICD\_DIAGNOSES) que definem os códigos usados no sistema.
- **Notas Clínicas:** NOTEEVENTS, que contém registros em texto livre, como sumários de alta e relatórios de radiologia.

## 4. Conteúdo Técnico

- **Formatos:** Os dados brutos são distribuídos em arquivos .csv.gz seguindo o padrão RFC 4180.
- **Dados Fisiológicos:** Inclui frequências cardíacas, pressão arterial, saturação de oxigênio e resultados de gasometria.
- **Informações Administrativas:** Códigos ICD-9 para diagnósticos e procedimentos, além de dados de faturamento.

## 5. Acesso e Ética

O acesso não é imediato; exige que o pesquisador complete o treinamento de ética em pesquisa com seres humanos (como o curso CITI Program) e assine um acordo de uso de dados no portal PhysioNet. (JOHNSON et al., 2016)

As informações contidas nesta base, incluem medições de sinais vitais, exames laboratoriais, escalas e índices aplicadas nos pacientes, procedimentos, tratamentos e notas clínicas de textos livres anonimizados. Ou seja, os pacientes não são identificados, preservando assim sua privacidade e cumprindo rigor também das leis brasileiras como a LGPD ([Lei Geral de Proteção de Dados](#)) (HOU et al., 2020).

Para acesso a esta base de dados, [MIMIC-III](#), algumas etapas, foram necessárias e

exigidas, como requisito essencial pelo detentor dos direitos de uso MIMIC-III - (MIT (*Massachusetts Institute of Technology*)).

**Há duas etapas obrigatórias que devem ser concluídas antes que o acesso a base seja concedido:**

- o pesquisador principal deve concluir um curso reconhecido na proteção de participantes de pesquisas humanas que inclui os requisitos da Lei de Portabilidade e Responsabilidade de Seguros de Saúde (HIPAA).
- O pesquisador deve assinar um acordo de uso de dados, que descreve o uso de dados apropriado e os padrões de segurança e proíbe esforços para identificar pacientes individuais.

A aprovação requer pelo menos uma semana. Uma vez que uma inscrição tenha sido aprovada, o pesquisador recebe por e-mails contendo instruções para baixar o banco de dados do *PhysioNetWorks*, um componente de acesso restrito do *PhysioNet*, uma espécie de repositório do MIT (*Massachusetts Institute of Technology*).

O Institucional Review Board do BIDMC concedeu uma renúncia ao consentimento informado e aprovou o compartilhamento do recurso de pesquisa. **(JOHNSON et al., 2016)**

Para organização do MIMIC-III, foram minerados em dois sistemas diferentes de informação de cuidados intensivos que estavam em vigor durante o período de coleta de dados: *Philips CareVue Clinical Information System (modelos M2331A e M1215A; Philips Healthcare, Andover, MA)* e *iMDsoft MetaVision ICU (iMDsoft, Needham, MA)*. **(JOHNSON et al., 2016 e HEAVNER et al., 2024)**

Vale destacar que, na fase 2.2 foi aplicado o método KDD (*Knowledge Discovery in Databases*), conforme ilustrado na **figura 19**, uma metodologia de tratamento de dados escolha para apoio a extração das características mais relevantes para a próxima fase. Nesta fase, optou-se por promover a mineração e tratamento dos dados estruturados, com técnicas de identificação de padrões, correlações e anomalias de registros. Este processo é essencial, pois permite estabelecer o processo de Descoberta de Conhecimento em Base de Dados, como sugere o método proposto por **(FAYYADD et al. 1996)**

**Figura 19 – Metodologia KDD (*Knowledge Discovery in Databases*)**



Fonte: Adaptação do modelo de (FAYYAD et al., 1996).

No curso da etapa 2.2, observou-se os arquivos do dataset e suas características extraídos e tratados, provenientes da base MIMIC-III:

**Tabela 2: Exibição dos DataSets Utilizados no Estudo**

		
Note_Eventos_Dados_Pacientes_c.csv	Processamento	Noteevents_textos_Outcome_c.csv
165 Mb		3,7 Gb
Corpus: 1.133.319 linhas		Corpus: 2.083.179 linhas

Fonte: Autor (2025)

### 3.3 LIMPEZA E PRÉ PROCESSAMENTO DA AMOSTRAGEM

Na **tabela 2**, onde exibe-se dois importantes *datasets*, chamados de “**Note\_Events-dados\_pacientes\_c.CSV**” e “**Noteevents\_textos\_Outcome\_c.CSV**” responsáveis pelo armazenamento de todo volume de dados estruturados e não estruturados, limpos, tratados com apoio da linguagem *Python* 3.0, no ambiente *COLAB PRO do Google*<sup>®</sup>, prontos para exibição, treinamento e processamento dos algoritmos de aprendizagem de máquina.

Estiveram elegíveis para processamento, 1.100.000 pacientes em correlação nos 2 *datasets*. A base de dados MIMIC-III é foi extraído os datasets utilizados na pesquisa, que gerou uma coleção de arquivos CSV (*comma-separated values*) com dados e valores separados por vírgula, juntamente com scripts para auxiliar na importação dos dados para sistemas de banco

de dados, incluindo *PostgreSQL*, *MySQL* e *MonetDB*. Como o banco de dados contém informações detalhadas sobre o atendimento clínico dos pacientes, ele deve ser tratado com o devido cuidado e é regido de leis internacionais de proteção de dados.

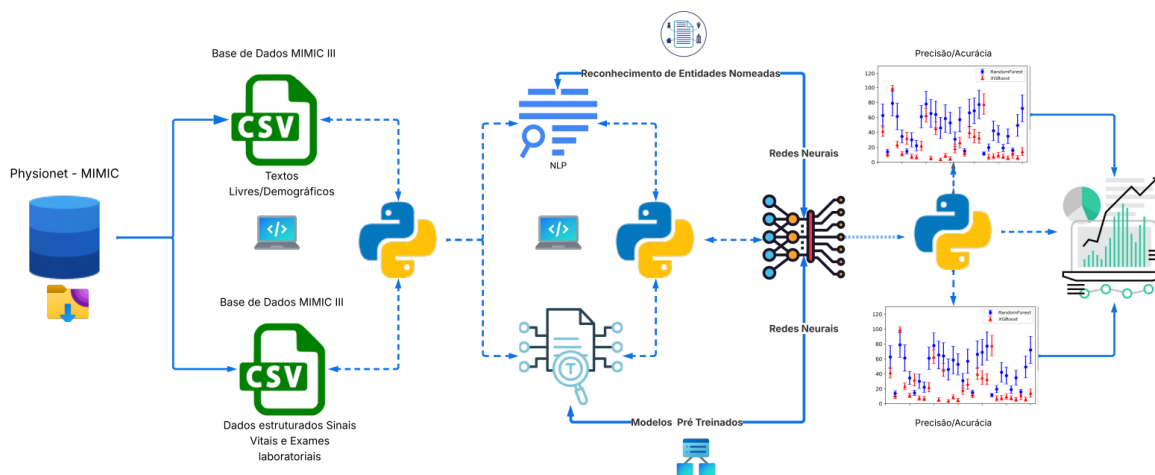
Esses dados utilizados na pesquisa são inicialmente reunidos de uma base do Microsoft SQL Server dentro de um *data warehouse* seguro do *Beth Israel Deaconess Medical Center* (BIDMC). Posteriormente, são transferidos para um sistema de banco de dados PostgreSQL (*PostgreSQL 12.8*, *PostgreSQL Global Development Group*) em servidores seguros do MIT por meio de uma conexão de Rede Privada Virtual (VPN). (MENG et al., 2021)

Os conjuntos de dados complementares, como definições de sistema de código (por exemplo, Classificação Internacional de Doenças (CID) e registros de óbitos estaduais são adquiridos independentemente do BIDMC e carregados no banco de dados PostgreSQL localizado no MIT (JOHNSON et al., 2023 e MENG et al., 2021)

De posse dos datasets, através do ambiente *COLAB PRO do Google*<sup>®</sup> com linguagem *Python 3.0*, foi iniciada a limpeza e pré-processamento dos dados; Elaborou-se os algoritmos de carregamento, leitura e normalização; Criou-se os modelos de treinamento e aprendizado supervisionado; Demonstrou-se os resultados e métricas de desempenho das técnicas de aprendizado e por fim, comparou-se os resultados e apresentamos as conclusões, limitações e sugestões sugeridas por. (HAQUE et al., 2022)

Na **figura 20**, observa-se uma ilustração visual para demonstração da origem dos algoritmos do índice proposto. Observa-se a procedência do MIMIC-III da *Physionet*, onde os dados foram limpos, tratados e organizados em dois *corpus* (1º *dataset* CSV com dados não estruturados com textos livres e o 2º *dataset* CSV com os dados estruturados, a partir do qual foi possível determinar o índice para a predição da deterioração clínica.

**Figura 20: Rede originária de processamento dos dados do índice proposto**



Fonte: Autor (2025)

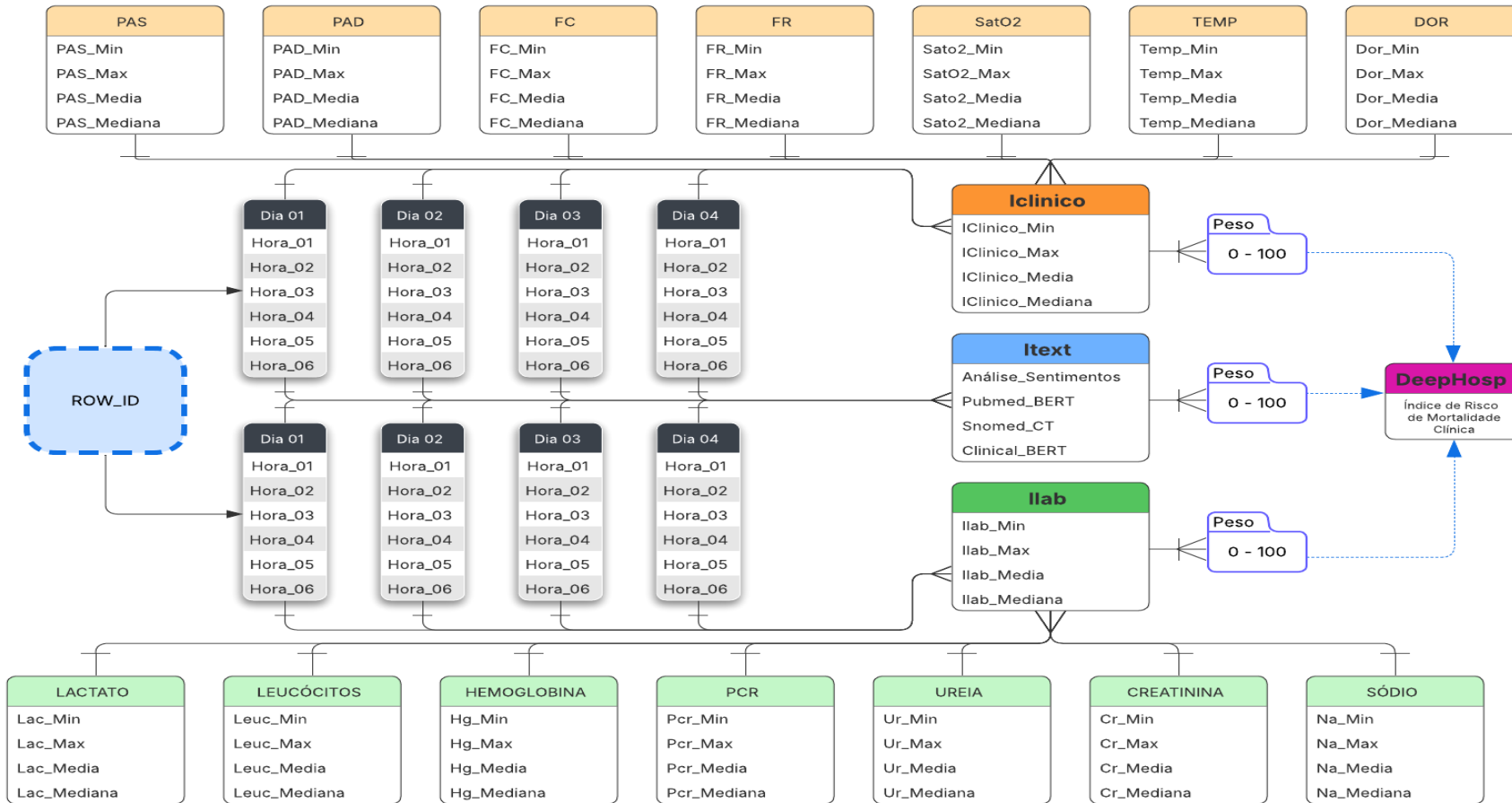
Assim, a presente pesquisa contempla as entradas (**Inputs**) de dados a partir das fontes: (1) Dados demográficos dos pacientes, (2) Registros dos sinais vitais, Escalas e índices aplicados ao paciente, (3) Exames laboratoriais e os (4) Registros de textos livres, como anotações e evoluções dos profissionais de saúde envolvidos na assistência ao paciente. Os inputs são direcionados a fase de limpeza e tratamento, normalizados, processados, treinados e por fim aplicados as métricas de acurácia e precisão. Portanto gerados os (**Outputs**) com os algoritmos treinados e demonstração do índice de riscos de deterioração dos pacientes.

### 3.3.1 ORIGEM DOS DADOS

Para se chegar ao índice de predição da condição clínica do paciente, conforme a proposta do estudo, empregou-se uma abordagem matemática que atribuiu pesos diferentes a sinais vitais e a exames laboratoriais, permitindo a quantificação de sua relevância preditiva. A seguir será detalhado como cada parcela do índice foi determinada.

Na **figura 21**, uma das principais representações da compreensão de origem dos dados de pesquisa. Os pacientes foram organizados no dataset, sendo no atributo “**ROW\_ID**” desidentificados. Apenas chamados de 01, 02, 03..... e assim prospectivamente. Outras variáveis numéricas para composição do grupo (atributo) que chamaremos de “**ICLÍNICO**” são conhecidas como: PAS = Pressão Arterial Sistólica, PAD = Pressão Arterial Diastólica, FC = Frequência Cardíaca, SatO<sup>2</sup> = Saturação de oxigênio no sangue, TEMP = Temperatura corporal e DOR = Score de dor. Outras variáveis numéricas para composição do grupo (atributo) que chamaremos de “**ILABORATÓRIO**” são conhecidas como: LACTATO, LEUCÓCITOS, HEMOGLOBINA, PCR, URÉIA, CREATINIA E SÓDIO. E outras variáveis com pacotes de modelos pré-treinados com princípios de NER (*Named Entity Recognition*), ou Reconhecimento de Entidades, como: ANÁLISE DE SENTIMENTOS, PUBMED\_BERT, SNOMED\_CT e CLINICAL\_BERT, como técnicas de Processamento de Linguagem natural para definição do grupo (atributo) que chamaremos de “**ITEXTO**”.

**Figura 21: Diagrama de exposição da origem dos dados e processamento do Índice**

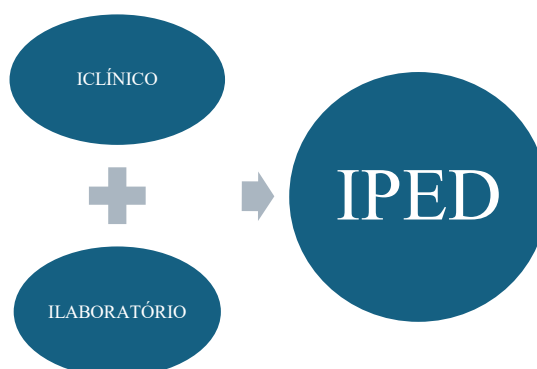


Fonte: Autor (2025)

### 3.4 DETERMINAÇÃO DO ÍNDICE DA PREDIÇÃO CLÍNICA DOS PACIENTES

Na **figura 22**, nota-se que para atingirmos **IPED**, **ICLÍNICO**, **ILABORATÓRIO** e **ITEXTO** foram somados, a fim de compor o indicador final, com um valor que variou entre 0 e 100%, a depender dos pesos atribuídos a qual cada variável de entrada.

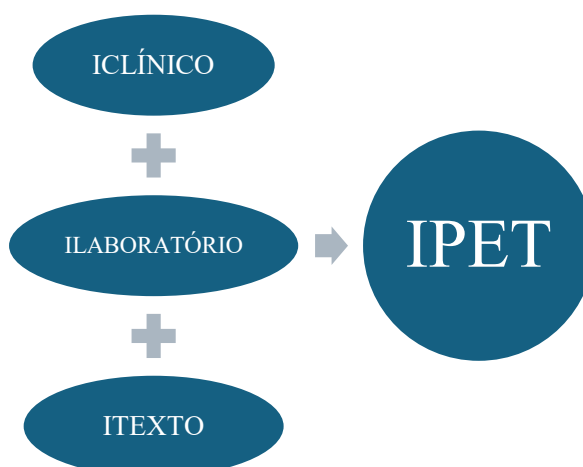
**Figura 22: Representação visual da junção de  
 $ICLÍNICO + ILABORATÓRIO = IPED$**



Fonte: Autor (2025)

Na **figura 23**, nota-se que para atingirmos **IPET**, **ICLÍNICO + ILABORATÓRIO + ITEXTO**, deverão somar juntos de a – 100 pontos a depender dos pesos atribuídos a qual cada variável de entrada.

**Figura 23: Representação visual da junção de  
 $ICLÍNICO + ILABORATÓRIO + ITEXTO = IPET$**



Fonte: Autor (2025)

### 3.5 PADRONIZAÇÃO DE CÁLCULO E EQUAÇÕES DOS MODELOS

Antes de calcular o índice, é preciso demonstrar como foi projetado a padronização de cálculo dos dados:

A **equação 3** coloca todas as variáveis na mesma escala (em torno do 0), evitar que uma com números extensos ou pequenos como o potássio. (SINGH et al, 2020)

#### Equação 3 – Modelo matemático para cálculo das variáveis

$$\mathbf{X\_padronizado} = (\mathbf{X} - \mathbf{média(X)}) / \mathbf{desvio-padrão(X)} \quad (3)$$

Fonte: Equação extraída de (SINGH, et al 2020)

Esta equação coloca todas as variáveis na mesma escala, isto é, para média 0 e desvio padrão unitário, de forma a evitar que a escala domine o modelo do indicador e permita ainda uma melhor otimização dos modelo de aprendizagem de máquina testados.

Para determinar o peso de cada parcela do IPED, mais particularmente de cada variável de ICLINICO e ILABORATÓRIO, aplicou-se uma regressão logística com regularização L1 (LASSO). Esse método é um tipo de estatística que escolhe quais variáveis realmente importam. Ele “zera” as que não ajudam, e oferece um peso maior e proporcional às que importam. A regularização L1, ou regressão lasso (“operador de seleção e encolhimento mínimo absoluto”), é um método de regularização que penaliza coeficientes de alto valor em um modelo de aprendizado de máquina. O peso de cada variável foi calculado por meio de uma fórmula adaptada de NING et al. (2024) para análise de importância de variáveis em modelos de *machine learning*.

Primeiro, aplica-se um modelo chamado **Regressão Logística com regularização L1**.

- **Regressão Logística** é um modelo estatístico que estima a probabilidade de um evento binário (por exemplo, alta = 1 ou óbito = 0) usando uma combinação linear das variáveis.
- Cada variável recebe um coeficiente  $B_j$ , que indica quanto ela contribui para aumentar ou diminuir a chance de alta.
- **Regularização L1** é uma técnica que força muitos coeficientes a ficarem exatamente iguais a zero, eliminando variáveis pouco relevantes e reduzindo risco de *overfitting*.

Depois do treinamento, o valor absoluto de cada coeficiente  $B_j$ , será dividido pela soma de todos os valores absolutos e obtemos um **peso relativo** para cada variável **equação 4**. A soma de todos os pesos é sempre 1 (ou 100%), o que permite interpretar cada peso como a **importância percentual** da variável dentro do grupo.

### 3.6 DEFINIÇÃO DO PESO DE CADA VARIÁVEL

**Equação 4 – Modelo matemático para cálculo e definição do peso de cada variável**

$$w_j = \frac{|\beta_j|}{\sum_k |\beta_k|}$$

(4) (6)

Fonte: Equação extraída de (TIBSHIRANI, R. 1996)

Assim, conforme a equação 5, aplica-se:  $\text{Peso\_de\_uma\_variável} = \text{valor\_absoluto\_de\_}B_j \div \text{soma\_dos\_valores\_absolutos\_de\_todos\_os\_}B_k$ . Cada variável ganha um peso proporcional à sua importância para prever o desfecho (alta ou óbito). Os pesos somam 100%. A seguir será detalhado como cada parcela do IPED, ou seja, o ICLINICO, o ILABORATÓRIO e o ITEXTO foram determinados.

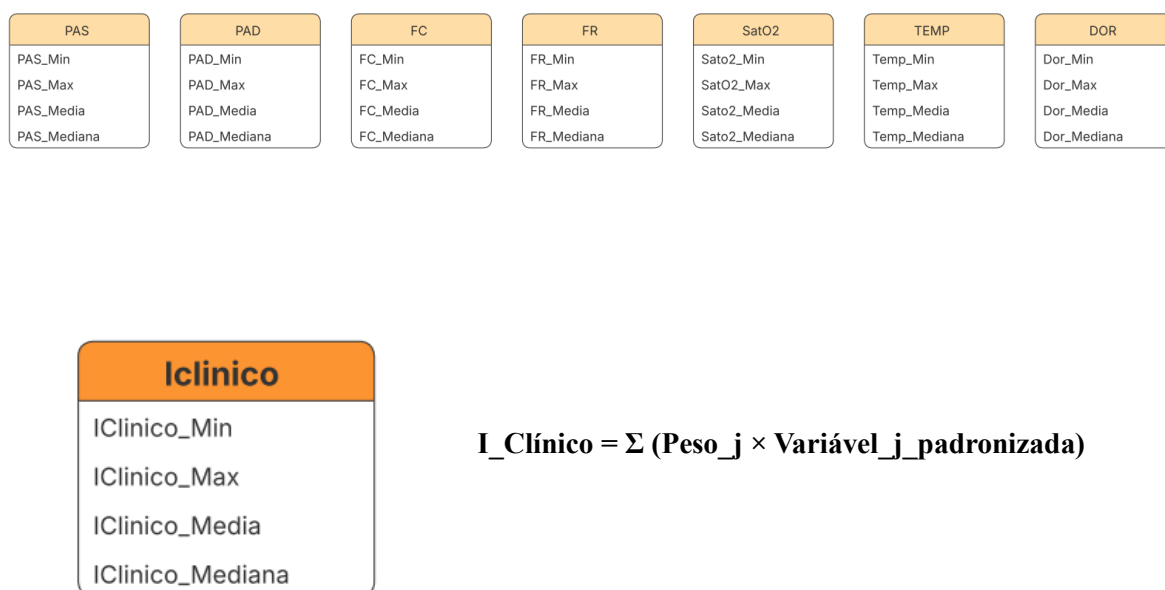
Para o cálculo do **IClínico**, foram observadas algumas variáveis como PAS (Pressão arterial sistólica), PAD (Pressão arterial diastólica), FC (Frequência cardíaca), FR (Frequência respiratória),  $\text{SatO}^2$  (Saturação de oxigênio no sangue), TEMP (Temperatura corporal) e DOR (Score de dor). Sabendo que em todas as variáveis, incorporados os valores Mínimos, Máximos, Médios e Medianos para serem submetidos a uma matriz de correlação e coeficiente de peso para cálculo do **IClínico**.

As variáveis determinadas e a equação apresentada na **figura 24** permitiram analisar as variáveis clínicas dos sinais vitais mais relevantes, conforme determinado por seus pesos, a fim de se determinar o valor final do **IClínico**. Assim, aplicou-se apenas as variáveis clínicas (temperatura, pressão, frequência cardíaca, etc.), multiplicou-se cada uma pelo seu peso e foi

somado o valor resultante. Assim temos um número único que resume o estado clínico, como sugere (SULLIVAN et al, 2004).

### 3.7 DEFINIÇÃO DO ÍNDICE ICLÍNICO

**Figura 24 – Variáveis dos sinais vitais e cálculo do IClínico**

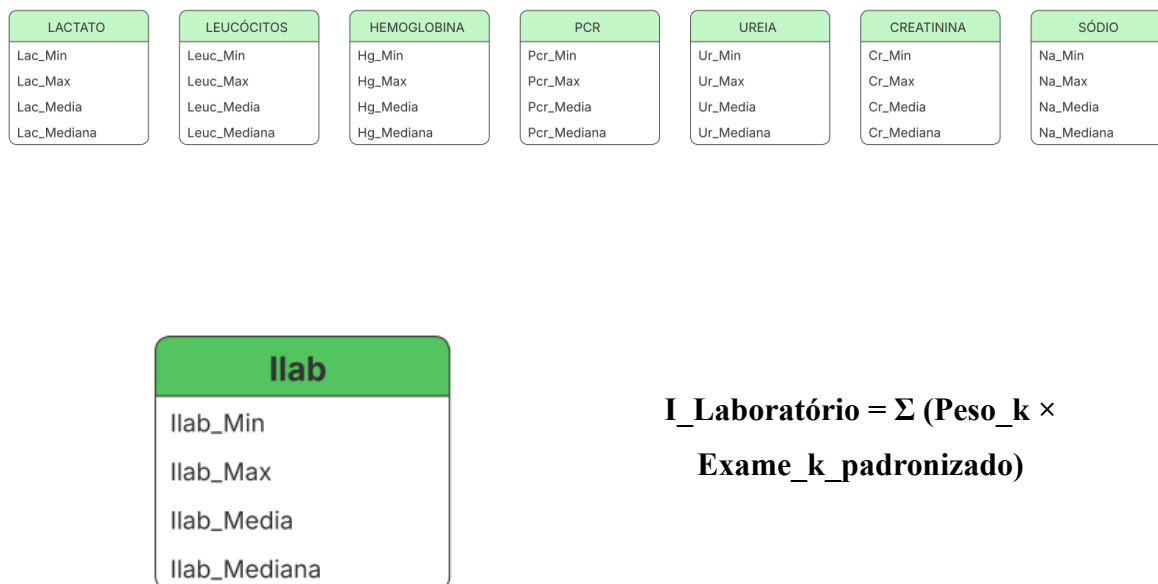


Fonte: Autor (2025)

De maneira análoga, foi determinada a parcela **ILaboratório na figura 25**. Neste caso, foram observados dados como LACTATO (Um produto do metabolismo da glicose em energia), LEUCÓCITOS (Leucócitos ou glóbulos brancos, são células produzidas pela medula óssea que tem a função de defender o organismo contra doenças, infecções e alergias. HEMOGLOBINA (Uma proteína encontrada nos glóbulos vermelhos do sangue, que tem como função transportar oxigênio dos pulmões para os tecidos do corpo). PCR (A proteína C-Reativa, é um marcador amplamente utilizado em análises laboratoriais para avaliar processos inflamatórios no estado de saúde dos pacientes). UREIA (É um marcador de saúde dos rins ou a condição da função renal do paciente). CREATININA (Também um marcador de saúde dos rins ou a condição da função renal do paciente). SÓDIO (É um dos principais, do grupo dos minerais essenciais para o organismo humano). Sabendo que em todas as variáveis, incorporados os valores Mínimos, Máximos, Médios e Medianos para serem submetidos a uma matriz de correlação e coeficiente de peso para cálculo do **ILaboratório**.

### 3.8 DEFINIÇÃO DO ÍNDICE ILABORATÓRIO

**Figura 25 – Variáveis dos exames de laboratório e cálculo de ILaboratório**



Fonte: Autor (2025)

Por fim, foram analisadas as anotações textuais obtidas de cada paciente da base MIMIC-III, a fim de calcular o ITexto. Neste caso, os textos foram tratados a partir de algumas ações resultantes de um processo de análise com técnicas de processamento de linguagem natural (PNL), com uso de modelos de transformers pré treinados com as chamadas Entidades Nomeadas, como: CLINICAL BERT, SNOMED CT, PUBMED BERT e com ANÁLISE DE SENTIMENTOS. De maneira similar aos cálculos de ILaboratorio e IClinico, a parcela ITexto foi calculada a partir de quatro grandezas geradas pelo processamento dos textos, incluindo as fases de reconhecimento das entidades mais relevantes e da análise de sentimentos, conforme mostrado na **figura 32**.

### 3.9 DEFINIÇÃO DO ÍNDICE ITEXTO

**Figura 26 – Variáveis dos resultados de análise dos textos livres, não estruturados com anotações e evoluções médicas e de enfermeiros**

Itext
Análise_Sentimentos
Pubmed_BERT
Snomed_CT
Clinical_BERT

$$I\_Texto = \Sigma (\text{Peso}_m \times \text{Variável\_de\_texto}_m)$$

Fonte: Autor (2025)

A equação demonstrado na **figura 26** permitiu o cálculo de análise das variáveis resultantes. Esse índice resume automaticamente o conteúdo textual em um número comparável com os demais índices como **ICLÍNICO** e **ILABORATÓRIO**.

Ao final, foi possível determinar o IPED, ou ainda o IPET, conforme demonstrado nas equações X1 e X2 a seguir.

### 3.10 DEFINIÇÃO DOS ÍNDICES COMBINADOS DE : IPED OU IPET

**Equações – Modelos matemáticos para cálculo de resultados de IPED e IPET**

$$IPED = \alpha.I_{Clínico} + \beta.I_{Laboratório}, \quad \alpha + \beta = 1 \quad (5)$$

$$IPET = \alpha.I_{Clínico} + \beta.I_{Laboratório} + \Delta.I_{Texto}, \quad \alpha + \beta + \Delta = 1 \quad (6)$$

Fonte: Adaptação ao modelo proposto por (NENES et al, 2009)

- O IPED será o índice que demonstra a condição do paciente apenas com Iclínico + Ilaboratorio. (**IPED = I\_Clínico + I\_Laboratório**).
- O IPET adiciona também o peso do texto não estruturado. (**IPET = I\_Clínico + I\_Laboratório + I\_Texto**)

O índice global IPET foi então calculado como uma **combinação linear adaptada de regressão logística**, conforme proposto por **DAS et al. (2025)** para avaliação de importância relativa de preditores. Finalmente, convertemos o IPET em probabilidade de alta ou óbito por meio de uma **função logística padrão**, conforme apresentado por (**HASTIE et al, 2009**).

### 3.11 DETALHAMENTO DA ETAPA 2.3 DA FASE 2

Como descrito no início deste capítulo, nesta fase foram construídos o banco de dados com os índices de milhares de pacientes e os algoritmos de aprendizagem de máquina para a predição temporal futura do índice de deterioração dos pacientes. Neste sentido, inicialmente foi construído um dataset com o IPED calculado para milhares de pacientes, em intervalos de 6 horas, como será melhor descrito no próximo capítulo.

Ainda na etapa 2.3, foi avaliado como tratar o problema da estimativa futura do índice IPED, ou IPET, a partir dos valores passados de cada paciente. Como se trata de uma série temporal, ou ainda de um problema sequencial, optou-se por empregar algoritmos de aprendizagem de máquina capazes de lidar com este tipo de problema, e no caso, foi escolhida uma arquitetura de rede neural de memória de longo e curto prazo (sigla LSTM em inglês) e uma rede convolucional (sigla CNN em inglês),

A escolha por integrar redes LSTM e CNN justifica-se em parte pela sua abordagem em outros estudos, como no trabalho de **TIAN, J., et al. (2024)**, mas também pela necessidade de capturar de forma robusta e eficiente as complexidades inerentes aos Registros Eletrônicos de Saúde (RES) para a previsão de riscos futuros de deterioração clínica e mortalidade. As redes LSTM são particularmente adequadas para modelar sequências temporais de dados clínicos, devido à sua capacidade de aprender dependências de longo prazo e de manter informações relevantes ao longo do tempo, mesmo em sequências de alta dimensionalidade e com padrões variáveis.

Por outro lado, as CNNs têm se mostrado altamente eficientes na extração de características locais e na identificação de padrões invariantes em dados estruturados ou semi-estruturados, o que é desejável na análise de sinais biológicos, exames laboratoriais, ou outras variáveis contínuas presentes nos RES. A aplicação de camadas convolucionais permite a captura de informações relevantes em diferentes escalas, facilitando a detecção de padrões sutis associados aos eventos adversos.

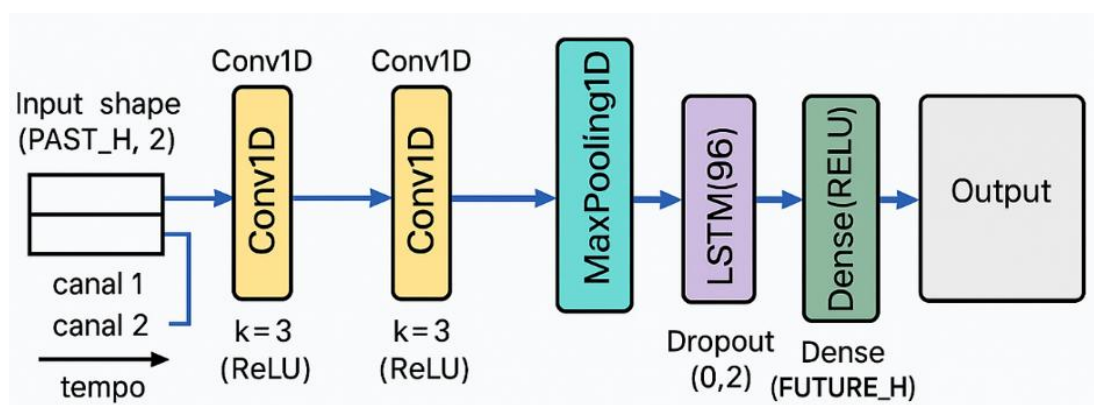
A combinação dessas arquiteturas resulta em um modelo híbrido que consegue explorar simultaneamente a dependência temporal de longo prazo e as características locais específicas

dos dados clínicos, potencializando a capacidade preditiva. Tal abordagem visa aprimorar a performance do sistema de previsão ao integrar o processamento sequencial e a extração de atributos discriminativos, contribuindo de forma significativa para o suporte à decisão clínica e a melhora dos desfechos de saúde.

Deste modo, a arquitetura de aprendizagem de máquina avaliada neste trabalho, apresentada na **figura 27**, consiste em um modelo híbrido que combina camadas de redes neurais convolucionais (CNN) e redes de memória de longo e curto prazo (LSTM). Esse tipo de arquitetura é frequentemente utilizado para processar dados de séries temporais, como dados de sensores ou sinais financeiros.

### 3.12 DEFINIÇÃO DA ARQUITETURA HÍBRIDA COM USO DE REDE NEURAL LSTM - CNNs

**Figura 27: Arquitetura de uma rede neural híbrida com uso de –CNN-LSTM**



Fonte: Autor (2025)

A seguir é apresentado um detalhamento das principais camadas e recursos que fazem parte da arquitetura escolhida:

- **Camadas Convolucionais (Conv1D):** A entrada é uma série temporal com duas características (canal 1 e canal 2) em um período passado (*PAST\_H*). As duas camadas *Conv1D* são usadas para extrair automaticamente padrões locais e características importantes da sequência de entrada. A convolução 1D é eficaz para identificar padrões espaciais ou temporais curtos nos dados. A função de ativação *ReLU* ajuda a introduzir não-linearidade no modelo.
- **Camada de Agrupamento (MaxPooling1D):** A camada *MaxPooling1D* reduz a dimensionalidade da saída das camadas convolucionais, mantendo as características mais importantes e diminuindo a quantidade de parâmetros a serem processados. Isso também ajuda a tornar o modelo mais robusto a pequenas variações nos dados de

entrada.

- **Camada Recorrente (LSTM):** A saída das camadas convolucionais é então passada para a camada LSTM(96). As LSTMs são um tipo de rede neural recorrente (RNN) que se destacam no processamento de dados sequenciais, pois são capazes de capturar dependências de longo prazo entre os intervalos de tempo. A camada *Dropout* é usada para prevenir o *overfitting* durante o treinamento.
- **Camadas Densas (Dense):** Após a camada LSTM, as camadas *Dense* (também conhecidas como camadas totalmente conectadas) são usadas para interpretar os recursos aprendidos e produzir a saída final. A última camada densa com *FUTURE\_H* neurônios provavelmente é usada para prever o valor da série temporal em um período de tempo futuro. Essa combinação de CNN e LSTM permite que o modelo capture simultaneamente padrões locais e dependências temporais de longo prazo nos dados.

### 3.13 DEFINIÇÃO DOS HIPERPARÂMETROS DA REDE NEURAL CNN – LSTM PROPOSTO PARA OS EXPERIMENTOS

**Tabela 3: Proposta de parametrização inicial para predição T +1 (6 horas)**

<b>SEED</b>	<b>PAST_H</b>	<b>FUTURE_H</b>
<b>TARGET_MODE</b>	<b>EPOCHS</b>	<b>BATCH_SIZE</b>
<b>VAL_SPLIT</b>		<b>LR</b>

1. **SEED = ?**
  - **O que é:** Um valor inicial para a "semente" de geração de números aleatórios.
2. **PAST\_H = ? (contexto)**
  - **O que é:** Número de passos de tempo anteriores considerados como contexto na entrada da rede (por exemplo, os últimos 22 valores de uma série temporal).
3. **FUTURE\_H = ? (horizonte futuro)**
  - **O que é:** Quantidade de passos à frente que o modelo deve prever.
4. **TARGET\_MODE = "level" ou "residual"**
  - **O que é:** Define o tipo de alvo que o modelo tenta aprender.
  - **Opções:**
    - **"level":** Prever a série original, normalizada ou não.

- **"residual"**: Prever a diferença (resíduo) entre o valor atual e um valor de referência.
5. **EPOCHS = ?**
    - **O que é:** Número de vezes que o modelo passará por todo o conjunto de treinamento durante o aprendizado.
  6. **BATCH\_SIZE = ?**
    - **O que é:** Número de exemplos processados de uma vez em uma única atualização dos pesos.
  7. **VAL\_SPLIT = ?**
    - **O que é:** Percentual dos dados reservados para validação durante o treinamento.
  8. **LR = 1 e -3 (taxa de aprendizado)**
    - **O que é:** Velocidade na qual os pesos do modelo são ajustados durante o treinamento.

No próximo capítulo serão apresentados os principais resultados obtidos com o uso dos algoritmos combinados CNN e LSTM para prever o valor futuro do IPED para os pacientes.

## 4.0 RESULTADOS E DISCUSSÃO

### 4.1 RESULTADOS PRELIMINARES DE PUBLICAÇÕES PRODUZIDAS NA PRIMEIRA FASE DA PESQUISA

Durante a aplicação do *checklist Extension for Scoping Reviews* (Checklist para revisão de escopo baseado em método PRISMA, onde o processo sistemático é usado como uma síntese de literatura integrativa para mapear a amplitude das evidências sobre o tópico, um conceito ou uma questão específica. Habitualmente a revisão de escopo é usada para uma etapa preliminar antes de uma revisão sistemática se necessário ou para identificar as lacunas do estado da arte existente. (HADDAWAY et al., 2022).

Importante destacar, que durante o cumprimento das disciplinas obrigatórias, ao longo das atividades desenvolvidas em grupo e com outros colegas e professores do programa de pós-graduação, foi possível o privilégio de publicar dois importantes trabalhos de pesquisa descritos a seguir no **quadro 14**.

#### Quadro14 – EXIBIÇÃO DAS PUBLICAÇÕES PRODUZIDAS DURANTE A ELABORAÇÃO DA TESE

Proposed use of data mining techniques in healthcare for knowledge generation in patient care: an integrative review.	Uso de aprendizado de máquina em registros eletrônicos de Saúde para prever deteriorações clínicas e mortalidade em pacientes adultos hospitalizados
 <p>ANÁLISE DO 11º SEMINÁRIO INTERNACIONAL CONGRESS OF HEALTH</p> <p>Proposed use of data mining techniques in healthcare for knowledge generation in patient care: an integrative review</p> <p><a href="https://doi.org/10.56238/homeIIsevenhealth-054">https://doi.org/10.56238/homeIIsevenhealth-054</a></p> <p>Reinaldo Ribeiro de Oliveira Sarah Loretta Barbosa dos Santos Renato Jose Sassi</p> <p>1 INTRODUCTION</p> <p>Data mining, also known as database knowledge discovery (KDD), is a process of extracting patterns, trends, and insights from large and complex data sets. It involves various techniques such as association rule mining, clustering, classification, and prediction, which can discover hidden patterns and relationships in data. Data mining has been widely used in various domains, including healthcare, to extract valuable insights that can inform decision making and improve outcomes (Pooja et al, 2019).</p> <p>In healthcare, data mining has gained significant attention due to the availability of large amounts of electronic health records (EHRs) and other health data, as well as the growing need for personalized, evidence-based care. The use of data mining techniques in healthcare has the potential to revolutionize patient care by providing insights into patient characteristics, disease patterns, treatment outcomes, and other relevant information. This can help healthcare professionals make informed decisions about diagnosis, treatment, and patient management, leading to better patient outcomes and improved healthcare delivery (N.P.J et al, 2022).</p> <p>On the current state of data mining in healthcare, Drayton-Brooks et al. (2020) suggests that data mining techniques have been increasingly used in healthcare to extract knowledge and insights from large and complex datasets such as electronic health records, clinical databases, administrative databases, and health insurance claims data.</p> <p>These datasets contain a wealth of information about patients, including demographics, medical history, diagnoses, treatments, laboratory results, and other relevant data that can be used for knowledge generation in patient care (Dreisbach C et al, 2019).</p> <p>The use of data mining in healthcare has shown promising results in several areas, such as disease prediction and risk assessment, treatment effectiveness evaluation, adverse event detection, patient stratification, and personalized medicine. For example, Ara Shaikh et al (2022) suggests that data mining techniques have been used to identify patterns and trends in patient data that can predict</p> <p>Anais do 11º Seven International Congress of Health</p>	 <p>REVISTA DE GESTÃO E SECRETARIADO</p> <p>ISSN: 2178-5010 DOI: <a href="http://doi.org/10.7769/gesec.v16i4.4882">http://doi.org/10.7769/gesec.v16i4.4882</a></p> <p>Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados</p> <p>Using Machine Learning in Electronic Health Records to Predict Clinical Deterioration and Mortality in Hospitalized Adult Patients</p> <p>Uso del Aprendizaje Automático en Registros Médicos Electrónicos para Predecir el Deterioro Clínico y la Mortalidad en Pacientes Adultos Hospitalizados</p> <p>Reinaldo Ribeiro de Oliveira<sup>1</sup> Cleber Gustavo Dias<sup>2</sup></p> <p>Resumo</p> <p>Introdução: O processo de hospitalização de pacientes envolve cuidados especializados, com equipes multiprofissionais de saúde, focados na recuperação e retomada do paciente ao estado de equilíbrio. Na hospitalização, pode acontecer deteriorações clínicas. Os Registros Eletrônicos de Saúde (RES) são, em sua forma mais simples, versões digitais dos prontuários dos pacientes. Eles disponibilizam informações instantaneamente, "quando e onde forem necessárias". Justificativa: Entendemos oportunidades de propor e produzir um experimento tecnológico computacional para condicionar um algoritmo matemático, aprender e prever mudanças de complicações de pacientes hospitalizados. Objetivo: Propor com aprendizado de máquina em registros eletrônicos de saúde, algoritmos de treinamento para prever deteriorações clínicas em pacientes adultos hospitalizados. Metodologia: A Fase1 primeira foi produzida uma revisão de escopo, na Fase2: foram construídos os experimentos com elaboração dos algoritmos de Machine Learning, onde utilizamos a base de dados MEDICIV com 233.452 pacientes com 20.000 registros tratados de pacientes para a pesquisa. Resultados: Foram produzidos 6 experimentos computacionais com treinamentos supervisionados, usando técnicas de Regressão Logística, Random Forest e XGBoost, cujos resultados demonstram</p> <p><sup>1</sup>Mestre em Enfermagem, Centro Universitário São Camilo, São Paulo, São Paulo, Brasil. <sup>2</sup>E-mail: <a href="mailto:mimada.gustavo.oliveira@gmail.com">mimada.gustavo.oliveira@gmail.com</a> (Brasil) <a href="https://orcid.org/0000-0002-1219-9862">https://orcid.org/0000-0002-1219-9862</a> <sup>3</sup>Pós-graduado em Engenharia Elétrica, Universidade de São Paulo, São Paulo, São Paulo, Brasil. E-mail: <a href="mailto:dias2019@gmail.com">dias2019@gmail.com</a> (Brasil) <a href="https://orcid.org/0000-0001-4212-2499">https://orcid.org/0000-0001-4212-2499</a></p> <p>Revista de Gestão e Secretariado – GeSec, V. 16, N. 4, P. 01-23, 2025 São José dos Pinhais, Paraná, Brasil.</p>
<p>Oliveira, R. R. de, Santos, S. L. B. dos, &amp; Sassi, R. J. (2023, August 14). Proposed use of data mining techniques in healthcare for knowledge generation in patient care: an integrative review. <i>11 seven international congress of health</i>. <a href="https://doi.org/10.56238/homeIIsevenhealth-054">https://doi.org/10.56238/homeIIsevenhealth-054</a></p>	<p>Oliveira, R. R. de, &amp; Dias, C. G. (2025). Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados. <i>Revista de Gestão e Secretariado</i>, 16(4), e4882. <a href="https://doi.org/10.7769/gesec.v16i4.4882">https://doi.org/10.7769/gesec.v16i4.4882</a></p>

Fonte: Autor (2025)

## 4.2 RESULTADOS OBTIDOS NA SEGUNDA FASE DA PESQUISA

Nesta fase foram conduzidos os diversos experimentos computacionais com o propósito de montar a base de dados dos indicadores calculados, bem como para o treinamento e testes com os algoritmos de aprendizagem de máquina avaliados. No caso, tais experimentos foram produzidos em um hardware do tipo Notebook LENOVO – *ThinkPad*, processador: 11th Gen Intel® Core™ i7-1165g7 @2.80 GHz (8 CPUs), memória de 8192 RAM Windows 11 pro 64 Bits (10.0, Compilação 261000), Disco Rígido de 1 Terabyte de armazenamento e placa de vídeo Nvidia Geforce Mx 110 com 2 Gigabytes de memória de vídeo. Os códigos criados nos experimentos, foram realizados também no ambiente, *COLAB PRO do Google*® com linguagem *Python* 3.0. Os softwares, bibliotecas e programas aplicativos usados nos experimentos são melhor descritos no anexo 3.

### 4.2.1 RESULTADOS OBTIDOS PARA A DETERMINAÇÃO DO IPED A PARTIR ( ICLÍNICO E ILABORATÓRIO)

Inicialmente, são apresentados os resultados gerados para o cálculo das variáveis mais relevantes. Na **figura 24**, observa-se a definição dos pesos de coeficiente das variáveis de **ICLÍNICO** utilizado com a técnica de regressão logística que correspondeu a 70,53% do valor total IPET. Na **figura 25**, observa-se a definição dos pesos de coeficiente das variáveis de **ILABORATÓRIO**, também obtido com a técnica de regressão logística, e que correspondeu no caso a 28,6% do IPET.

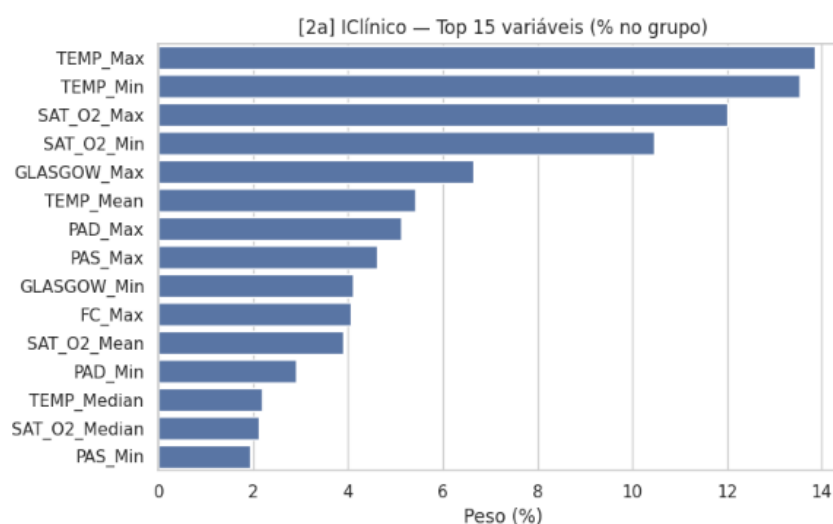
Na **figura 28**, observa-se a definição dos pesos de coeficiente das variáveis de **ICLÍNICO** utilizado com a técnica de regressão logística que logo ocupará o espaço de 70,53% do IPET.

### 4.3 CATEGORIZAÇÃO DOS PESOS DOS COEFICIENTES DAS VARIÁVEIS DE ICLÍNICO

**Figura 28 – Resultados dos achados de peso das variáveis no ICLÍNICO**

```

=== [2a] Top 30 variáveis - Iclínico (% no grupo) ===
TEMP_Max      13.85
TEMP_Min      13.54
SAT_O2_Max    12.01
SAT_O2_Min    10.45
GLASGOW_Max   6.64
TEMP_Mean     5.41
PAD_Max       5.12
PAS_Max       4.62
GLASGOW_Min   4.12
FC_Max        4.06
SAT_O2_Mean   3.89
PAD_Min       2.89
TEMP_Median   2.18
SAT_O2_Median 2.11
PAS_Min       1.94
PAD_Median    1.47
PAS_Mean      0.89
DOR_Median    0.83
FR_Min        0.83
GLASGOW_Mean 0.81
FC_Min        0.74
FC_Median     0.41
FR_Max        0.39
DOR_Min       0.34
FR_Mean       0.21
DOR_Max       0.20
FR_Median     0.03
PAD_Mean      0.00
GLASGOW_Median 0.00
FC_Mean       0.00
  
```



Fonte: Autor (2025)

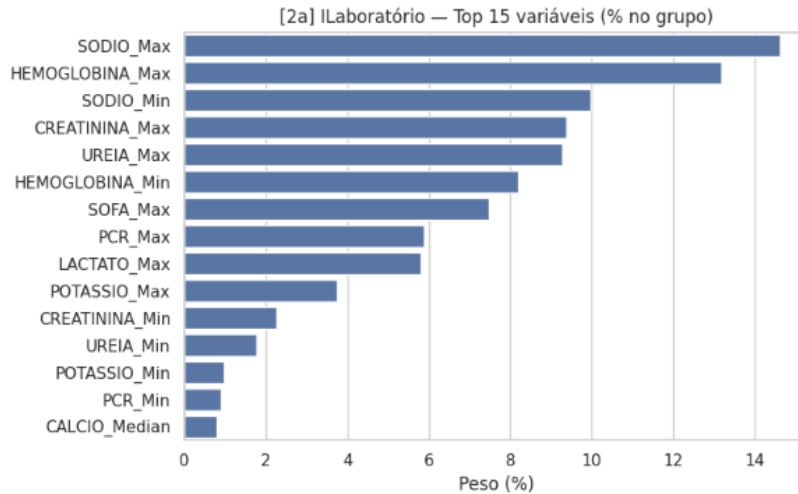
Na **figura 29**, observa-se a definição dos pesos de coeficiente das variáveis de **ILABORATÓRIO** utilizado com técnica de regressão logística que logo ocupara o espaço de 28,6% do IPET.

#### 4.4 CATEGORIZAÇÃO DOS PESOS DOS COEFICIENTES DAS VARIÁVEIS DE ILABORATÓRIO

**Figura 29 – Resultados dos achados de peso das variáveis no ILABORATÓRIO**

```

=== [2a] Top 40 variáveis - ILaboratório (% no grupo) ===
SODIO_Max      14.61
HEMOGLOBINA_Max 13.17
SODIO_Min      9.97
CREATININA_Max 9.37
UREIA_Max     9.28
HEMOGLOBINA_Min 8.18
SOFA_Max      7.47
PCR_Max       5.88
LACTATO_Max   5.80
POTASSIO_Max  3.75
CREATININA_Min 2.26
UREIA_Min     1.77
POTASSIO_Min  0.96
PCR_Min       0.90
CALCIO_Median 0.80
LACTATO_Min   0.78
LEUCOCITOS_Median 0.69
SODIO_Mean    0.65
POTASSIO_Mean 0.53
LEUCOCITOS_Min 0.45
CREATININA_Median 0.39
CALCIO_Min    0.36
LEUCOCITOS_Max 0.33
HEMOGLOBINA_Median 0.24
SOFA_Min     0.20
SOFA_Median   0.18
BICARBONATO_Max 0.14
UREIA_Median  0.14
PCR_Median    0.12
HEMATOCRITO_Min 0.11
LACTATO_Mean  0.10
HEMATOCRITO_Median 0.10
HEMATOCRITO_Max 0.08
BICARBONATO_Min 0.06
PCR_Mean      0.05
CALCIO_Max    0.05
BICARBONATO_Mean 0.03
HEMATOCRITO_Mean 0.02
LACTATO_Median 0.01
BICARBONATO_Median 0.01
    
```

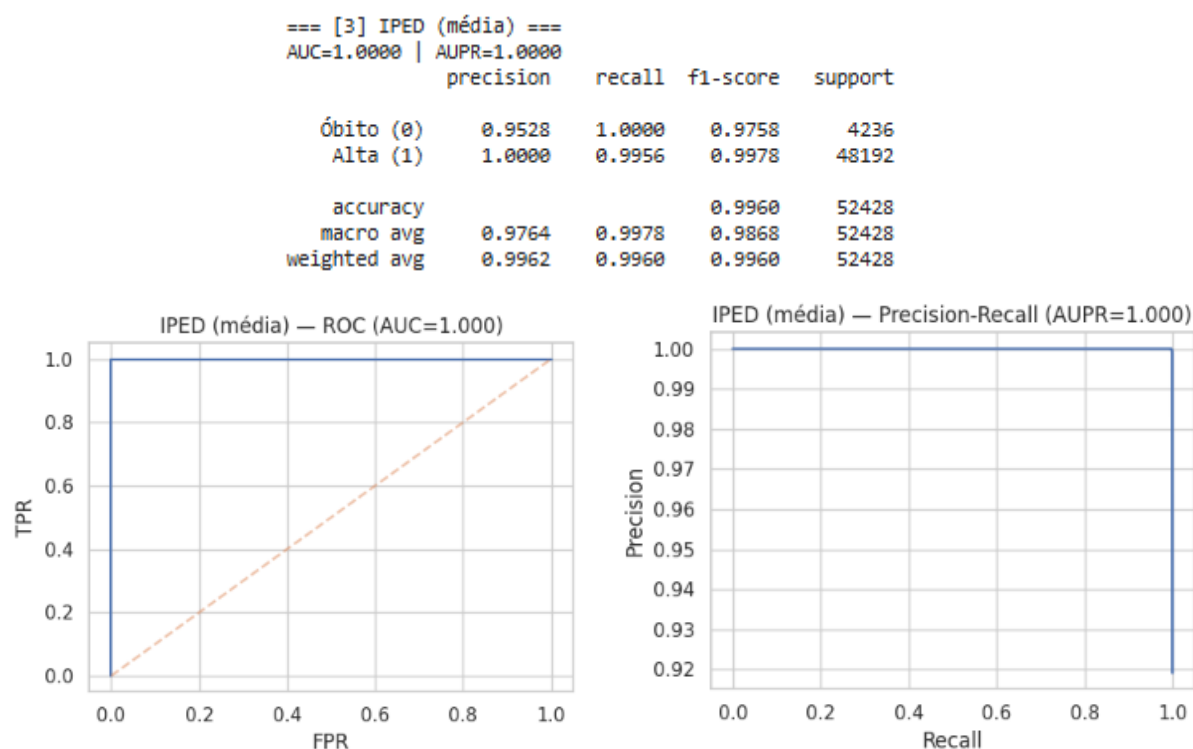


Fonte: Autor (2025)

Na **figura 30**, pode-se observar as métricas de desempenho da análise de IPED – Média, sabendo que o modelo identificou uma AUC de 100%, Uma Precisão de 95% para Óbitos e 100% para Altas, Um Recall para Óbitos de 100% e para Altas de 99%, Um F1-score de 97% para Óbitos e 99% para Altas, com um (N = 54.428)

#### 4.5 MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED - MÉDIA

**Figura 30 – Métricas de desempenho para o cálculo de IPED - MÉDIA**

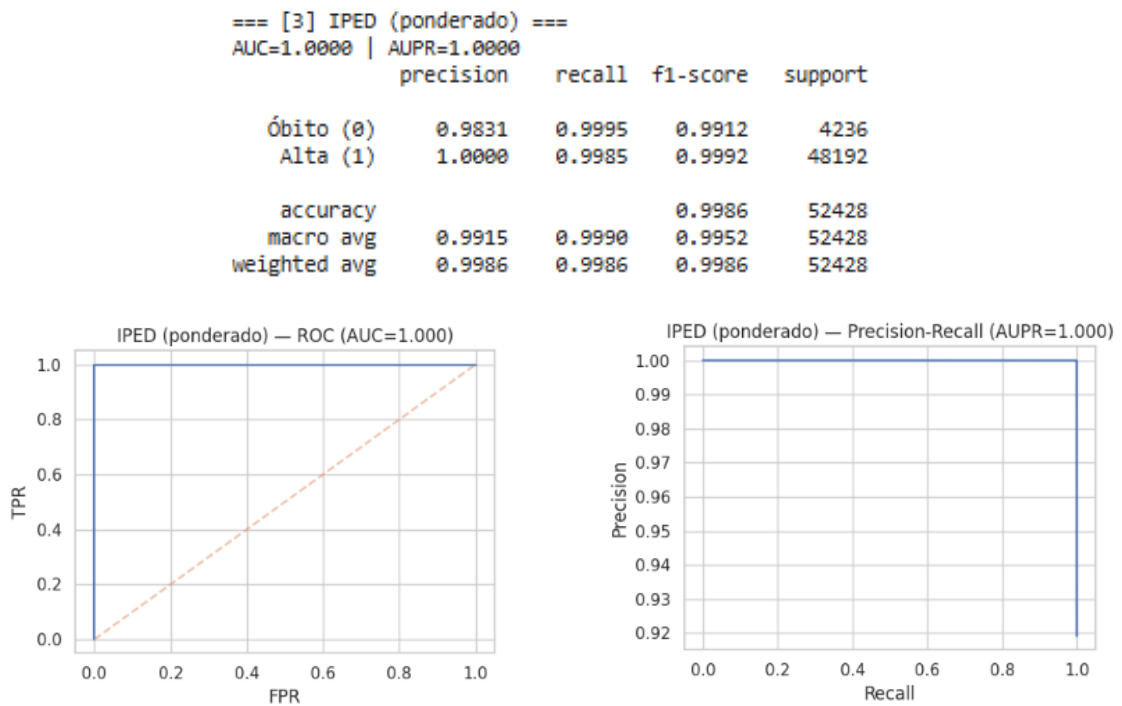


Fonte: Autor (2025)

#### 4.6 MÉTRICAS DE DESEMPENHO PARA O CÁLCULO DE IPED - PONDERADO

Na **figura 31** pode-se observar as métricas de desempenho da análise de IPED Ponderado, sabendo que o modelo identificou uma AUC de 100%, Uma Precisão de 98% para Óbitos e 100% para Altas, Um Recall para Óbitos de 99% e para Altas de 99%, Um F1-score de 99% para Óbitos e 99% para Altas, com um (N = 54.428).

**Figura 31 – Métricas de desempenho para o cálculo de IPED - PONDERADO**

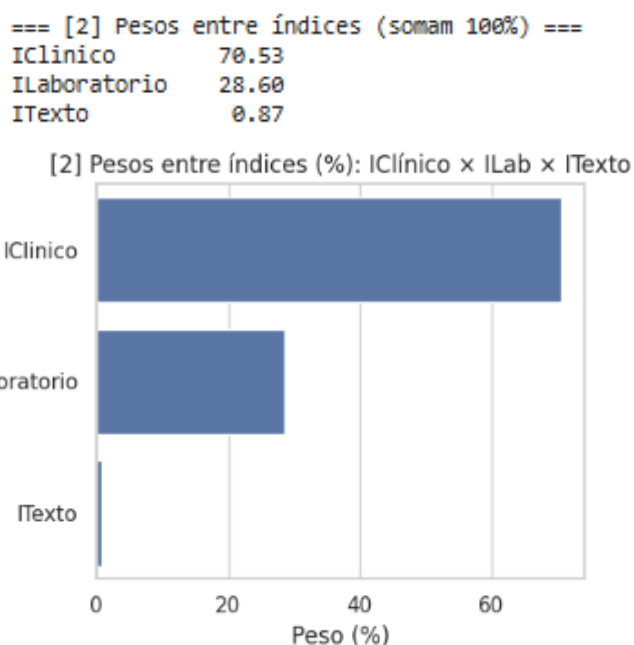


Fonte: Autor (2025)

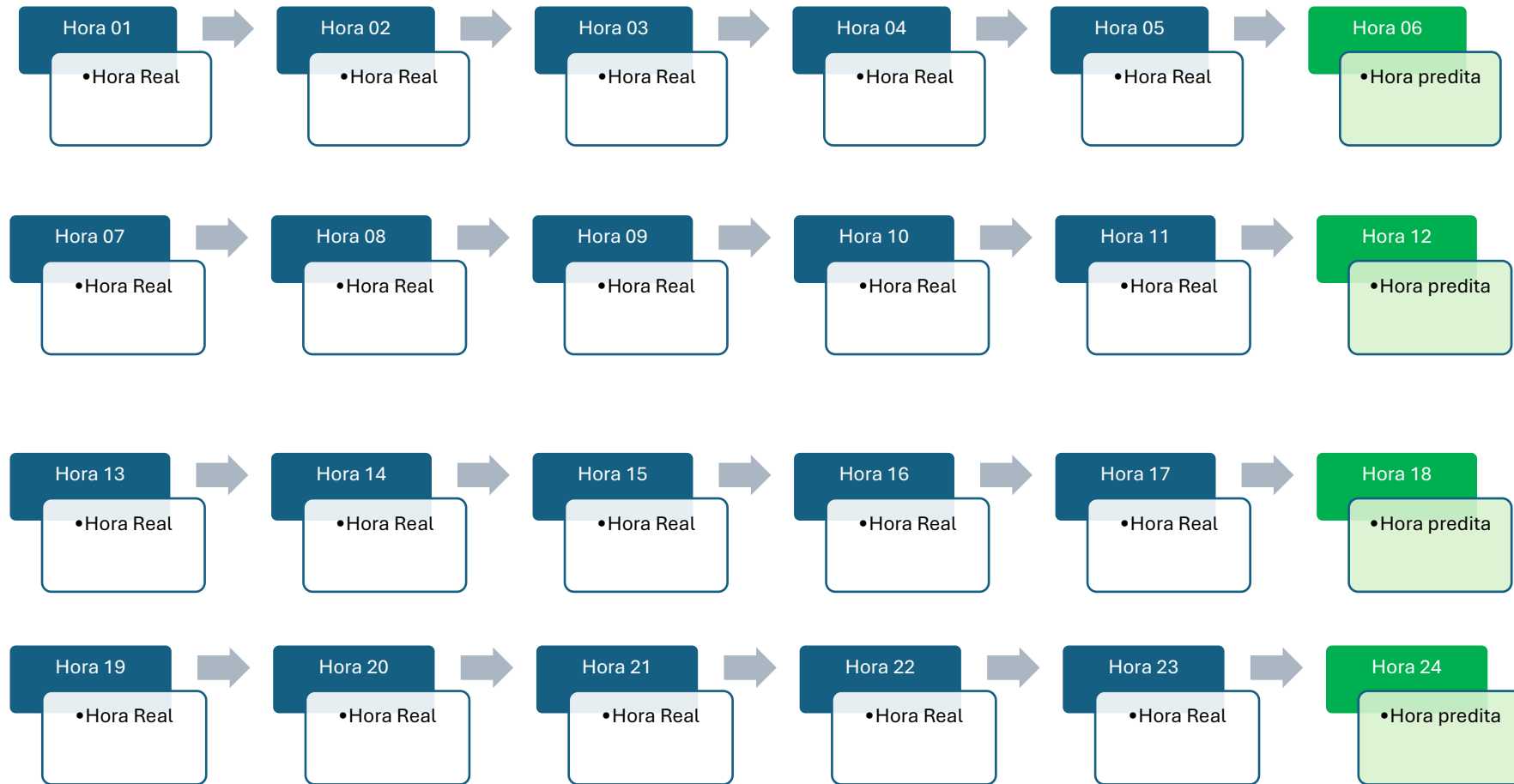
Na **figura 32**, exibe-se as métricas estabelecidas finais obtidos para o do peso de cada parcela, onde  $ICLÍNICO = 70,53\%$  do peso de IPET,  $ILABORATÓRIO = 28,60\%$  do peso de IPET e que  $ITEXTO = 0,87\%$  do peso de IPET. Como o valor do ITEXTO foi inferior a 1% do total, a sua parcela foi desconsiderada, sendo mantido portanto, para fins de predição, apenas o IPED.

#### 4.7 MÉTRICAS IDENTIFICADAS NO ACUMULADO DE PESO DO CÁLCULO DE IPET

**Figura 32 – Métricas identificadas no acumulado de peso do cálculo de IPET**



Fonte: Autor (2025)

**Figura 33 – Projeção ilustrativa da proposta de pesquisa**

Fonte: Autor (2025)

Na **figura 33** projeta-se a proposta visual do objetivo dos algoritmos de aprendizado de máquina

**Figura 34 – Exibição amostral do dataset, gerado a partir dos dados reais com os cálculos de IPED**

Row_ID	Hora01	Hora02	Hora03	Hora04	Hora05	Hora06	Hora07	Hora08	Hora09	Hora10	Hora11	Hora12	Hora13	Hora14	Hora15	Hora16	Hora17	Hora18	Hora19	Hora20	Hora21	Hora22	Hora23	Hora24
7	40.980564	75.278358	48.443691	48.246964	61.767895	60.310375	61.759102	38.429871	42.344387	39.032661	39.058208	38.354038	35.186855	52.009453	34.693928	41.488602	50.924156	68.453262	52.659016	35.175720	27.992920	24.898232	35.217148	32.552265
15	38.437748	37.680435	36.659023	60.033077	39.502262	28.827286	41.349022	29.738276	66.955765	63.845715	68.311775	33.201607	68.264351	69.800560	45.371307	64.670837	63.597298	44.256763	79.467308	51.142235	30.661476	69.194427	69.828667	46.824181
17	54.348034	42.475094	49.410912	32.003368	35.650650	63.819794	25.064556	45.761814	44.091576	32.228104	34.400505	32.766800	51.728771	9.724392	55.738644	45.475876	50.746719	50.309540	34.639347	24.402508	40.126766	44.365536	57.746700	37.606300
19	31.923815	33.938038	63.837833	31.647110	37.249161	29.241777	33.642036	45.223354	46.568897	60.960129	57.970985	53.192398	42.218658	17.294577	57.652477	38.453907	43.501747	39.995457	33.744762	27.124704	70.636345	44.818531	38.057117	35.515144
22	63.102585	37.079132	68.638466	39.101074	52.540104	49.698254	58.610592	33.121590	49.671368	53.980278	48.892044	73.646606	42.475887	26.018999	46.201775	38.046192	22.582109	28.101967	58.564930	51.426662	28.094582	46.966778	50.099621	32.639656
23	46.582264	35.400433	36.567513	44.732964	48.760303	51.743942	32.952892	22.382668	50.891022	58.721405	46.337994	32.407494	31.977758	38.925423	38.404072	51.009857	34.549374	35.437702	32.647182	33.800293	24.962505	29.337517	49.227394	31.581940
28	19.822094	43.616745	50.976032	59.763927	65.438438	35.194942	64.854912	40.860313	41.784138	47.056179	47.429932	44.498058	30.367100	45.573322	55.660263	32.040169	52.793640	36.432491	34.106308	57.427601	39.303150	46.429466	52.491570	45.339550
36	60.695549	56.621773	49.832287	54.717636	28.214045	42.383842	63.249222	33.265472	50.240223	21.913246	57.777042	63.798096	53.747250	31.463261	46.226852	66.059113	56.645809	26.830599	54.372360	44.207100	49.406956	24.623508	44.511719	
45	35.493599	62.019058	29.936625	38.623711	31.895170	70.472710	44.945976	34.169189	38.323284	46.881073	38.594589	28.654448	57.511127	68.780785	65.609131	59.029091	46.249676	40.821358	35.602325	37.238708	47.050877	58.894886	46.695518	27.469559
46	33.015934	67.743439	56.335461	47.395756	44.665146	67.633926	44.596966	22.156778	44.308331	64.838348	28.950356	55.131531	55.251865	30.786476	32.404995	30.445257	54.813148	32.282730	48.852535	64.667770	51.364182	47.076443	66.384895	68.658043
47	79.368126	52.559090	58.234726	74.324944	59.918758	30.341616	56.922737	41.635994	80.555389	51.454853	45.273930	78.228081	67.494476	77.494308	64.625412	29.224056	44.642864	63.631382	73.054398	36.578407	62.516899	49.752205	70.714157	51.027794
58	48.426620	40.659557	23.484663	66.014771	42.901581	24.226210	43.609608	61.667793	40.331043	61.581173	51.600769	51.568863	49.951946	40.220829	34.434532	23.209629	65.330727	43.736645	67.835747	31.642233	34.442478	27.711084	33.122627	33.965069
61	34.858002	64.760239	32.786259	55.240620	64.226799	34.717819	25.728258	61.648624	33.720066	29.349930	52.333786	45.677307	39.220852	42.829964	57.630840	13.957448	46.838837	31.422779	41.595634	35.866074	66.804878	51.401134	56.889359	32.962570
66	25.020243	57.251205	56.737915	39.275414	52.410522	40.159725	39.983861	32.274529	36.715759	65.233124	63.806700	56.327576	63.786484	39.129417	32.569424	41.878166	53.054344	53.821011	51.327068	46.422756	50.195004	33.341312	34.605377	53.223084
67	54.971218	35.035313	29.967260	37.540302	63.404934	18.300362	62.654884	55.057678	48.235043	38.132669	65.107887	73.057526	35.166098	38.793839	50.347027	56.506516	52.707371	34.286994	33.206444	22.577911	45.179000	22.577948	63.621165	47.427098
71	56.945648	45.385082	37.633392	47.268688	39.307095	61.156086	49.515644	22.637552	54.310005	43.244270	47.320572	33.217587	54.762482	50.642929	67.502785	36.930199	36.716549	34.258827	46.138832	43.715733	49.075001	52.176647	61.130829	78.435997
79	20.785236	42.533504	74.724854	41.786465	36.328320	50.926603	65.750595	48.021584	42.671875	36.596218	49.090477	42.226845	38.625618	52.415905	31.481686	37.128422	38.252220	47.475231	38.878963	51.619617	22.078863	35.983776	38.335052	48.785118
80	34.836246	51.324120	65.131004	50.675137	35.495567	56.153564	43.093864	53.026058	52.219921	34.419796	66.003220	64.766533	64.875626	61.227509	30.094433	47.792263	18.540262	57.940987	59.109798	32.041451	25.361958	65.373390	32.423458	11.743265
87	45.343498	67.068008	49.643749	42.079247	33.744152	56.559052	39.661339	28.215298	35.616184	32.742630	46.281157	49.828346	35.692978	46.484680	74.892502	71.949715	40.785362	33.269417	44.981071	30.853130	33.018623	30.568001	60.997200	40.578167
89	49.301769	23.128866	51.061676	61.632381	74.712029	62.659321	36.690464	34.420692	50.691315	58.583897	59.332146	59.691147	38.843712	35.893757	60.732437	60.247849	25.797632	57.612839	70.820778	60.703373	55.317520	53.358204	59.194183	
93	66.693977	55.056320	71.518105	58.710171	30.999746	40.248096	24.423580	64.641953	51.895870	28.561396	19.058100	42.511261	40.965416	49.244411	46.082855	39.448116	44.138138	42.995186	54.859005	51.395027	19.672531	57.220013	33.032925	29.720964
97	41.156548	36.061314	30.813513	53.271404	41.429657	50.156433	26.947016	43.473732	33.636478	47.444584	47.909393	50.583862	62.551285	51.228931	35.623039	68.210381	58.591900	36.295181	75.572403	18.481413	32.161068	36.324615	45.183456	43.950333
105	67.641838	55.209648	48.076996	81.222763	63.411793	69.103867	39.907623	70.036293	46.880959	33.404087	54.268322	65.257530	74.704483	72.893509	31.718086	69.357552	66.914711	61.945477	49.063385	57.206482	42.829651	69.560829	63.760563	69.906075
110	57.310455	62.682983	26.396593	47.443344	20.219263	37.019060	53.127102	44.629265	27.337795	47.316616	33.493484	63.080399	42.783157	38.633888	51.709045	46.811176	61.182225	45.213768	48.862587	48.344074	58.941719	47.047329	47.240368	45.303329
123	40.529713	37.062981	21.911213	39.629662	48.151840	59.433056	37.641560	45.556240	24.415670	68.358177	23.658503	51.933441	52.358479	47.883595	46.711414	31.975670	37.473618	23.655609	42.425911	55.827488	48.153313	46.830128	63.320221	45.367157
130	47.222542	32.045914	51.837116	45.038357	63.329189	39.478542	34.288403	37.024174	63.069706	56.763069	60.064232	41.177517	52.959721	57.052170	40.092468	65.837730	52.447407	25.338133	27.076767	38.645973	67.006142	45.758976	49.885620	66.963882
151	56.094563	64.247520	19.736122	65.352974	53.154579	72.482910	19.299114	37.281467	59.828602	37.909004	24.354603	30.390566	25.453447	57.135162	33.377220	53.163631	51.925190	37.227211	33.285683	28.822544	51.828259	61.192184	39.538971	52.466412
152	54.232452	34.463276	40.863720	62.045723	37.813507	47.921444	43.751362	53.139748	60.238430	70.053032	31.842789	41.862148	47.230766	50.353802	35.943432	29.688585	59.959442	40.642742	40.939434	35.542282	49.215858	53.760197	30.984533	45.397385
157	17.171280	52.406998	52.890553	68.328423	44.076302	52.046265	39.823235	52.403046	36.695641	46.744881	33.464700	38.192059	59.678841	47.437016	31.040577	14.896108	59.488461	47.388641	32.572018	29.429455	59.832291	44.687321	64.337494	67.726830
164	60.453506	64.550919	35.324806	46.326763	28.337725	50.336643	32.709587	37.186668	25.665454	23.653154	28.154886	24.338404	39.495548	43.258171	59.637733	41.706985	55.204578	36.116951	33.938679	46.992924	23.156342	59.990105	71.140244	35.247719
166	70.345764	54.282368	17.106020	34.502174	29.909010	36.422726	32.230286	26.276707	60.884148	53.389816	38.412903	33.190422	43.532585	29.702734	14.676206	39.125416	60.396763	27.551588	52.633411	65.594162	67.757103	65.223389	23.237047	67.541443
169	60.604259	49.562523	46.707367	47.461582	54.189980	34.573174	56.317688	38.300728	47.882790	66.313576	57.859386	45.740467	56.957684	58.646172	70.936989	45.633038	51.328213	69.522110	29.149866	21.632568	41.721375	25.771212	43.668530	46.627632
173	60.6552																							



## **4.9 PROPOSTAS DOS CENÁRIOS TESTADOS NOS EXPERIMENTOS COMPUTACIONAIS**

### **1-) Previsão usando como alvo do treinamento o resíduo (delta) – 2 canais**

Casos: t+1, t+2, t+3 e t+4

### **2-) Previsão usando o IPED diretamente (*sem resíduo*)**

Casos: t+1, t+2, t+3 e t+4

### **3-) Previsão usando 1 canal como alvo do treinamento o resíduo**

CASOS: t+1 e t+2

### **4-) Teste da rede usada no cenário 1, treinada com o dataset MIMIC-III, a partir dos dados disponibilizados pelo dataset MIMIC-IV.**

Casos: t+1 e t+2



No primeiro cenário da **figura 36**, observa-se os seguintes resultados:

#### **Distribuição das Faixas de MAPE**

- **MAPE inferior a 5%:** 655 pacientes
  - Indica que a maioria das previsões (655 de 1000) teve erro percentual absoluto abaixo de 5%, demonstrando alta precisão.
- **$5\% \leq \text{MAPE} < 10\%$ :** 230 pacientes
- **$10\% \leq \text{MAPE} < 15\%$ :** 73 pacientes
- **$15\% \leq \text{MAPE} \leq 20\%$ :** 28 pacientes
- **MAPE superior a 20%:** 14 pacientes
  - Poucas previsões apresentaram erro percentual superior a 20%, indicando que o modelo é bastante preciso na maioria dos casos.

#### **Métricas Globais**

- **Horizonte de previsão (h=1):** 1 passo adiante
- **MAE (Erro Médio Absoluto):** 2.0223
  - Em média, a previsão do modelo diverge em aproximadamente 2 unidades na escala da variável prevista, indicando a magnitude média do erro.
- **MAPE (Erro Percentual Médio Absoluto):** 4.9262%
  - Em média, as previsões se desviam cerca de 5% do valor real, reforçando a alta precisão do modelo para esse horizonte de previsão.

#### **Considerações Técnicas**

- A distribuição detalhada dos MAPE indica que a maior parte das previsões é altamente precisa, com erro abaixo de 5%. Isso demonstra a eficácia do modelo em capturar as variações do sistema.
- O MAE de aproximadamente 2 unidades sugere uma boa aproximação na escala dos dados, enquanto o MAPE abaixo de 5% reforça a confiabilidade das previsões.
- O uso de uma janela de histórico de 22 passos e um horizonte de previsão curto (1 passo) têm sido eficazes, refletindo nos resultados altamente precisos.

## 4.12 PROPOSTAS DOS CENÁRIOS TESTADOS - 2

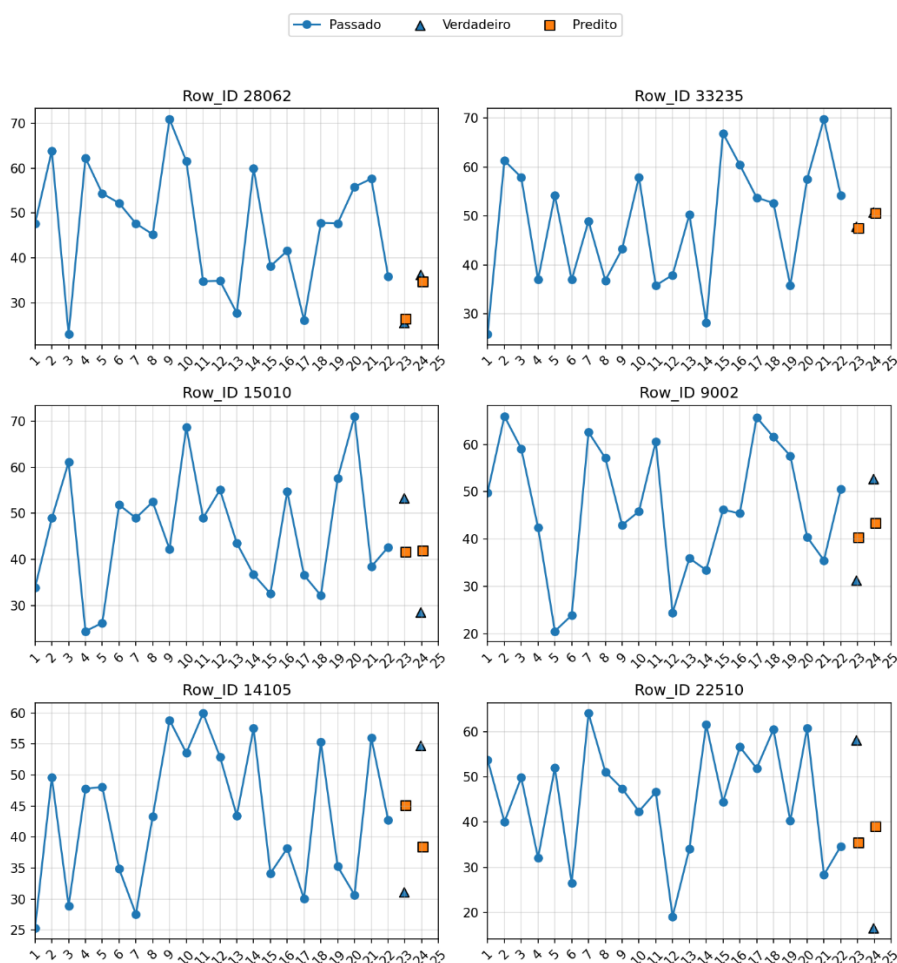
### Previsão da Rede CNN – LSTM, usando como alvo do treinamento o resíduo (delta) 2 canais

#### 2-) PREVENDO T+2 (Próximas 12 horas)

**Figura 37: Previsão: Usando alvo do treinamento e o resíduo (delta) predição T + 2 (12 horas)**

Foram utilizados 7.000 pacientes para Treinamento e Validação e 1.000 pacientes para testes

SEED = 42  
VAL\_SPLIT = 0.10  
PAST\_H = 22  
EPOCHS = 200  
FUTURE\_H = 2  
BATCH\_SIZE = 128



#### Métricas de Desempenho

=== Contagem por faixas de MAPE (TESTE) ===

< 5% : 183  
5% ≤ MAPE < 10% : 173  
10% ≤ MAPE < 15%: 148  
15% ≤ MAPE ≤ 20%: 115  
> 20% : 381

=== MÉTRICAS GLOBAIS (TESTE) ===

MAE : 7.6297

MAPE : 18.9822 %

Fonte: Autor (2025)

No segundo cenário da **figura 37**, observa-se os seguintes resultados:

### **Distribuição das Faixas de MAPE**

- **MAPE < 5%:** 183 pacientes
- **5% ≤ MAPE < 10%:** 173 pacientes
- **10% ≤ MAPE < 15%:** 148 pacientes
- **15% ≤ MAPE ≤ 20%:** 115 pacientes
- **MAPE > 20%:** 381 pacientes

A maior quantidade de previsões se concentra na faixa de erro acima de 20%, refletindo que uma parcela significativa das previsões teve erro relativo elevado. Ainda assim, há uma quantidade considerável de previsões com alta precisão (abaixo de 10%), indicando que o modelo é capaz de fazer previsões com bom desempenho em muitos casos.

### **Métricas Globais**

- **MAE (Erro Médio Absoluto):** 7.6297
  - Em média, as previsões do modelo apresentam um erro absoluto de aproximadamente 7,63 unidades na escala dos dados.
- **MAPE (Erro Percentual Médio Absoluto):** 18.9822%
  - Em média, as previsões divergem cerca de 19% dos valores reais, evidenciando uma margem de erro moderada em relação ao valor total previsto.

### **Considerações Técnicas**

- A distribuição das faixas de MAPE demonstra uma dispersão significativa no desempenho, com uma grande quantidade de previsões apresentando erro acima de 20%. Isso pode indicar a presença de outliers, variabilidade nos dados ou dificuldade na previsão em certas condições.
- Apesar disso, há uma proporção significativa de previsões com erro inferior a 10%, indicando que o modelo consegue captar bem o comportamento geral do sistema para uma parcela importante dos casos.
- O MAE de aproximadamente 7,63 sugere uma média de erro que, dependendo da escala dos dados, pode ou não ser aceitável. A métrica MAPE indica que, em média, há uma discrepância de cerca de 19%, o que pode ser considerado moderado ou alto, dependendo do contexto clínico ou de aplicação.
- O horizonte de 2 passos (FUTURE\_H = 2) traz um desafio maior para a precisão, pois prever múltiplos passos à frente normalmente reduz a acurácia.



No terceiro cenário da **figura 38**, observa-se os seguintes resultados:

#### **Distribuição das Faixas de MAPE**

- **MAPE < 5%:** 18 pacientes
  - Poucas previsões tiveram erro muito baixo, indicando alta precisão em alguns casos.
- **5% ≤ MAPE < 10%:** 108 pacientes
- **10% ≤ MAPE < 15%:** 161 pacientes
- **15% ≤ MAPE ≤ 20%:** 179 pacientes
- **MAPE > 20%:** 534 pacientes
  - A maioria das previsões (quase 54%) apresentou erro percentual superior a 20%, mostrando que a maior parte das previsões apresenta maior dispersão.

#### **Métricas Globais**

- **MAE (Erro Médio Absoluto):** 9.3636
  - Em média, o erro absoluto na previsão é de aproximadamente 9,36 unidades na escala dos dados.
- **MAPE (Erro Percentual Médio Absoluto):** 24.1958%
  - Considerando a média dos erros relativos, o modelo apresenta uma margem de erro de cerca de 24%, o que indica uma previsão moderadamente precisa, porém com dispersão considerável.

#### **Considerações Técnicas**

- A distribuição dos erros mostra que o maior volume de previsões apresenta erro superior a 20%, indicando dificuldades na precisão em várias instâncias, possivelmente devido à variabilidade dos dados ou a limitações do modelo em capturar padrões complexos num horizonte de 3 passos.
- O MAE de aproximadamente 9,36 unidades, combinado com um MAPE de quase 24%, sugere que a precisão geral pode variar bastante dependendo do padrão ou do paciente, sendo resultados adequados em alguns cenários e menos em outros.
- O horizonte de previsão de 3 passos (FUTURE\_H=3) é relativamente desafiador, especialmente com uma janela de histórico de 20, contribuindo para a dispersão dos erros.
- Esses resultados indicam potencial para ajustes, como otimizações de hiperparâmetros ou aumento de dados, para melhorar a estabilidade e a precisão do modelo.



No quarto cenário da **figura 39**, observa-se os seguintes resultados:

### Resultados no Teste

Distribuição das Faixas de MAPE (Erro Percentual Absoluto Médio)

- **< 5%:** 6 pacientes
  - Muito poucas previsões tiveram erro extremamente baixo, indicando alta precisão em uma minoria dos casos.
- **5% ≤ MAPE < 10%:** 53 pacientes
- **10% ≤ MAPE < 15%:** 140 pacientes
- **15% ≤ MAPE ≤ 20%:** 157 pacientes
- **> 20%:** 644 pacientes
  - A maioria das previsões (mais de dois terços) apresentou erro superior a 20%, o que sugere dificuldades na previsão para uma grande proporção de casos.

Métricas Globais

- **MAE (Erro Médio Absoluto):** 9.3636
  - A média do erro absoluto nas previsões foi de aproximadamente 9,36 unidades, refletindo a dispersão dos valores previstos em relação aos valores reais.
- **MAPE (Erro Percentual Médio Absoluto):** 24.1958%
  - Em média, as previsões diferiram cerca de 24,2% dos valores reais, indicando um nível moderado de precisão geral, embora com alta dispersão.

### Considerações Técnicas

- A grande quantidade de previsões com erro acima de 20% sugere dificuldades em capturar todos os padrões dos dados, possivelmente devido à complexidade do problema ou à variação de alguns casos específicos.
- A presença de uma pequena fração de previsões com erro inferior a 10% demonstra que, em certos cenários, o modelo consegue fornecer previsões muito próximas do valor real.
- O horizonte de previsão de 4 passos ( $FUTURE\_H = 4$ ) aumenta a complexidade da tarefa, refletida na dispersão dos erros, embora o modelo ainda mantenha uma média razoável de erro absoluto.
- Para melhorar o desempenho, estratégias como ajuste de hiperparâmetros, incremento de quantidade de dados, ou técnicas de ensemble podem ser considerados.



No quinto cenário da **figura 40**, observa-se os seguintes resultados:

### Resultados no Teste

#### Distribuição das Faixas de MAPE

- **< 5%:** 700 pacientes
  - Um número expressivo de previsões muito precisas, com erro abaixo de 5%, representando uma alta performance em diversos casos.
- **5% ≤ MAPE < 10%:** 227 pacientes
- **10% ≤ MAPE < 15%:** 47 pacientes
- **15% ≤ MAPE ≤ 20%:** 18 pacientes
- **> 20%:** Apenas 8 pacientes
  - A maioria das previsões teve erro muito baixo, demonstrando que o modelo é capaz de fornecer previsões altamente precisas na maior parte dos casos, com apenas uma pequena fração apresentando erro elevado.

#### Métricas Globais

- **MAE (Erro Médio Absoluto):** 1.7391
  - Em média, a previsão diverge aproximadamente 1,74 unidades do valor real, indicando alta precisão.
- **MAPE (Erro Percentual Médio Absoluto):** 4.2762%
  - Em média, as previsões diferem cerca de 4,3% dos valores reais, reforçando a efetividade do modelo na previsão de uma variável contínua com alta acurácia.

#### Considerações Técnicas

- A grande maioria das previsões apresenta erro abaixo de 10%, com uma parcela significativa tendo erro inferior a 5%, o que evidencia a alta precisão do modelo para prever o próximo valor na sequência.
- O baixo valor médio de MAE (1.74 unidades) combinado com um MAPE de aproximadamente 4,3% demonstra que o modelo é eficiente na captura do comportamento esperado, especialmente para previsões de nível imediato.
- A previsão de um passo à frente (FUTURE\_H = 1) mostra-se particularmente eficaz, como refletido na alta quantidade de previsões com erro muito baixo.
- Limitações podem estar relacionadas à variação de casos específicos, mas os resultados indicam uma performance robusta para aplicações em nível de previsão de curto prazo.



No sexto cenário da **figura 41**, observa-se os seguintes resultados:

### Resultados no Teste

#### Distribuição das Faixas de MAPE (Erro Percentual Absoluto Médio)

- **< 5%:** 166 pacientes
- **5% ≤ MAPE < 10%:** 190 pacientes
- **10% ≤ MAPE < 15%:** 131 pacientes
- **15% ≤ MAPE ≤ 20%:** 139 pacientes
- **> 20%:** 374 pacientes

Observa-se que uma parcela significativa das previsões apresentou erro abaixo de 10%, com uma quantidade importante também abaixo de 15%. Contudo, ainda há uma quantidade expressiva de previsões com erro superior a 20%, indicando que o desempenho varia bastante entre os casos.

#### Métricas globais

- **MAE (Erro Médio Absoluto):** 7.5579 unidades
  - Em média, a previsão difere cerca de 7,56 unidades do valor real, refletindo um grau de precisão razoável para previsão de nível.
- **MAPE (Erro Percentual Médio Absoluto):** 18.8760%
  - O erro médio relativo às previsões é de aproximadamente 19%, indicando uma margem moderada de erro, embora com dispersão considerável.

#### Considerações Técnicas

- Os resultados mostram que o modelo é capaz de fazer previsões bastante precisas em uma parte significativa dos casos, com muitas previsões apresentando erro abaixo de 10%.
- Ainda há uma quantidade considerável de previsões com erro alto (>20%), sugerindo potencial para melhorias, como ajuste de hiperparâmetros, aumento de volume de dados ou implementação de técnicas de otimização adicionais.
- A previsão de dois passos à frente (FUTURE\_H=2) revela-se desafiadora, refletida na dispersão dos erros, porém o desempenho geral ainda é satisfatório para aplicações cujo limite de tolerância seja próximo de 19%.



No sétimo cenário da **figura 42**, observa-se os seguintes resultados:

### **Resultados de Desempenho no Teste**

#### **Distribuição das faixas de MAPE**

- **< 5%:** 20 pacientes
  - Poucas previsões tiveram erro extremamente baixo, refletindo que o modelo consegue alta precisão em um número pequeno de casos.
- **5% ≤ MAPE < 10%:** 114 pacientes
- **10% ≤ MAPE < 15%:** 145 pacientes
- **15% ≤ MAPE ≤ 20%:** 168 pacientes
- **> 20%:** 553 pacientes
  - A maioria das previsões apresentou erro superior a 20%, indicando dificuldade de precisão em uma parcela expressiva dos casos no horizonte de 3 passos à frente.

#### **Métricas globais**

- **MAE (Erro Médio Absoluto):** 9.4255
  - Em média, as previsões se desviam aproximadamente 9,43 unidades do valor real, indicando uma margem de erro moderada a alta para o nível de previsão.
- **MAPE (Erro Percentual Médio Absoluto):** 24.6809%
  - A média do erro relativo é de aproximadamente 25%, o que sugere que há uma variação significativa na acurácia das previsões, com muitas tendo erro maior que 20%.

#### **Considerações Técnicas**

- Embora a maior parte das previsões esteja concentrada na faixa de erro acima de 20%, há uma parcela considerável com erro relativamente baixo, o que demonstra potencial de previsibilidade.
- O horizonte de 3 passos aumenta a complexidade da tarefa e, por isso, a dispersão dos erros é maior.
- Melhorias na modelagem, como ajustes nos hiperparâmetros, aumento de volume de dados ou técnicas de ensemble, podem ajudar a reduzir o erro global e aumentar a precisão em diferentes casos.

## 4.18 PROPOSTAS DOS CENÁRIOS TESTADOS – 8

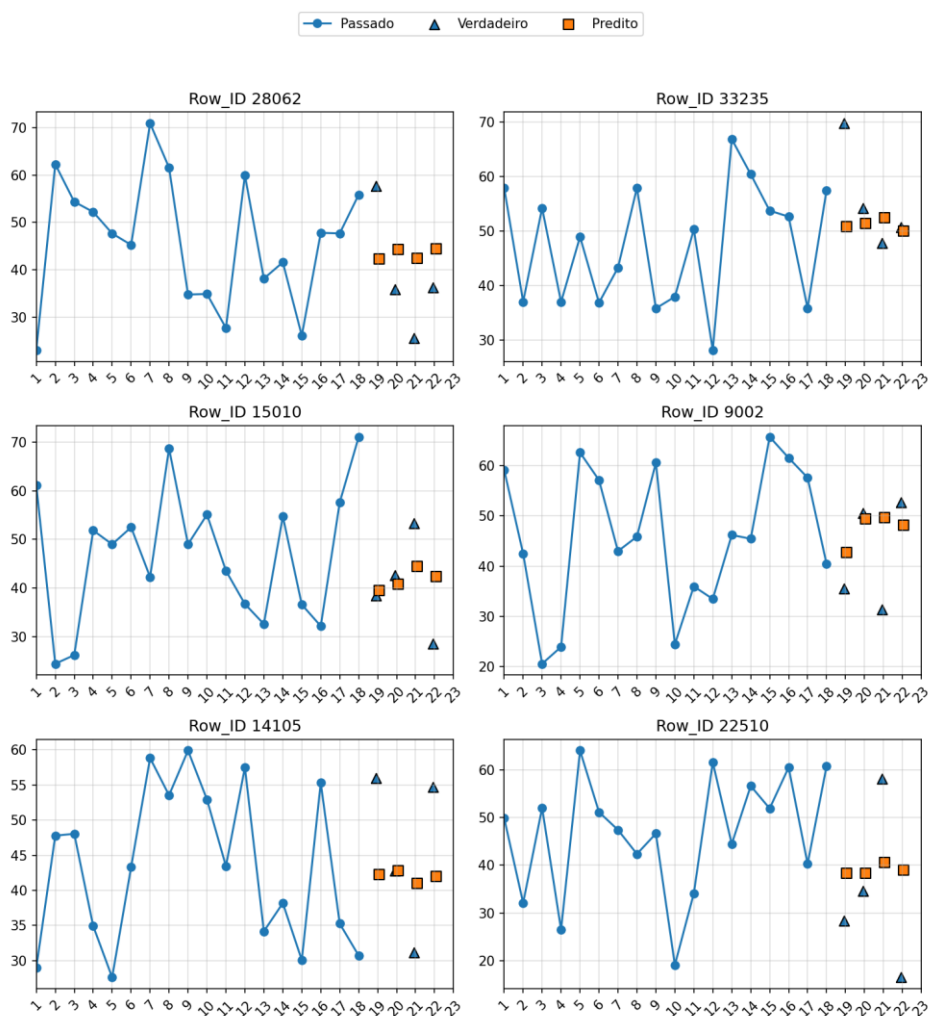
### Previsão da Rede CNN – LSTM, usando o IPED diretamente com 2 canais

#### 4-) PREVENDO T + 4 (Próximas 24 horas)

**Figura 43: Previsão: Usando como alvo do treinamento sem resíduo, previsão T + 4 (24 horas)**

Foram utilizados 7.000 pacientes para Treinamento e Validação e 1.000 pacientes para testes

SEED = 42  
VAL\_SPLIT = 0.10  
PAST\_H = 18  
EPOCHS = 200  
FUTURE\_H = 4  
BATCH\_SIZE = 128



#### Métricas de Desempenho

=== Contagem por faixas de MAPE (TESTE) ===

< 5% : 3  
5% ≤ MAPE < 10% : 53  
10% ≤ MAPE < 15% : 144  
15% ≤ MAPE ≤ 20% : 164  
> 20% : 636

=== MÉTRICAS GLOBAIS (TESTE) — previsão de NÍVEL ===

MAE : 9.8981  
MAPE : 25.9536 %

Fonte: Autor (2025)

No oitavo cenário da **figura 43**, observa-se os seguintes resultados:

### **Resultados no teste**

#### **Distribuição das faixas de MAPE (Erro Percentual Médio Absoluto)**

- **< 5%:** 3 pacientes
  - Pouquíssimas previsões extremamente precisas, indicando que o modelo consegue altitude de precisão em poucos casos.
- **5% ≤ MAPE < 10%:** 53 pacientes
- **10% ≤ MAPE < 15%:** 144 pacientes
- **15% ≤ MAPE ≤ 20%:** 164 pacientes
- **> 20%:** 636 pacientes
  - A maioria das previsões apresentou erro superior a 20%, refletindo dificuldades na previsão de uma grande parcela dos casos no horizonte de 4 passos.

#### **Métricas globais**

- **MAE (Erro Médio Absoluto):** 9,8981
  - Em média, o erro absoluto na previsão é aproximadamente 9,90 unidades, indicando que há uma dispersão média moderada.
- **MAPE (Erro Percentual Médio Absoluto):** 25,9536%
  - O erro médio relativo às previsões é de aproximadamente 26%, sugerindo uma margem de erro considerável na maioria dos casos, especialmente em previsão de múltiplos passos à frente.

#### **Considerações Técnicas**

Apesar de o modelo alcançar previsões altamente precisas em uma pequena fração de casos (< 5%), a maior parte dos resultados apresenta erro moderado a elevado, especialmente ao prever múltiplos passos à frente. Para melhorar a acurácia, recomenda-se ajustes nos hiperparâmetros, aumento de dados ou implementação de técnicas avançadas de modelagem, buscando reduzir o impacto da dispersão dos erros.

## 4.19 PROPOSTAS DOS CENÁRIOS TESTADOS – 9

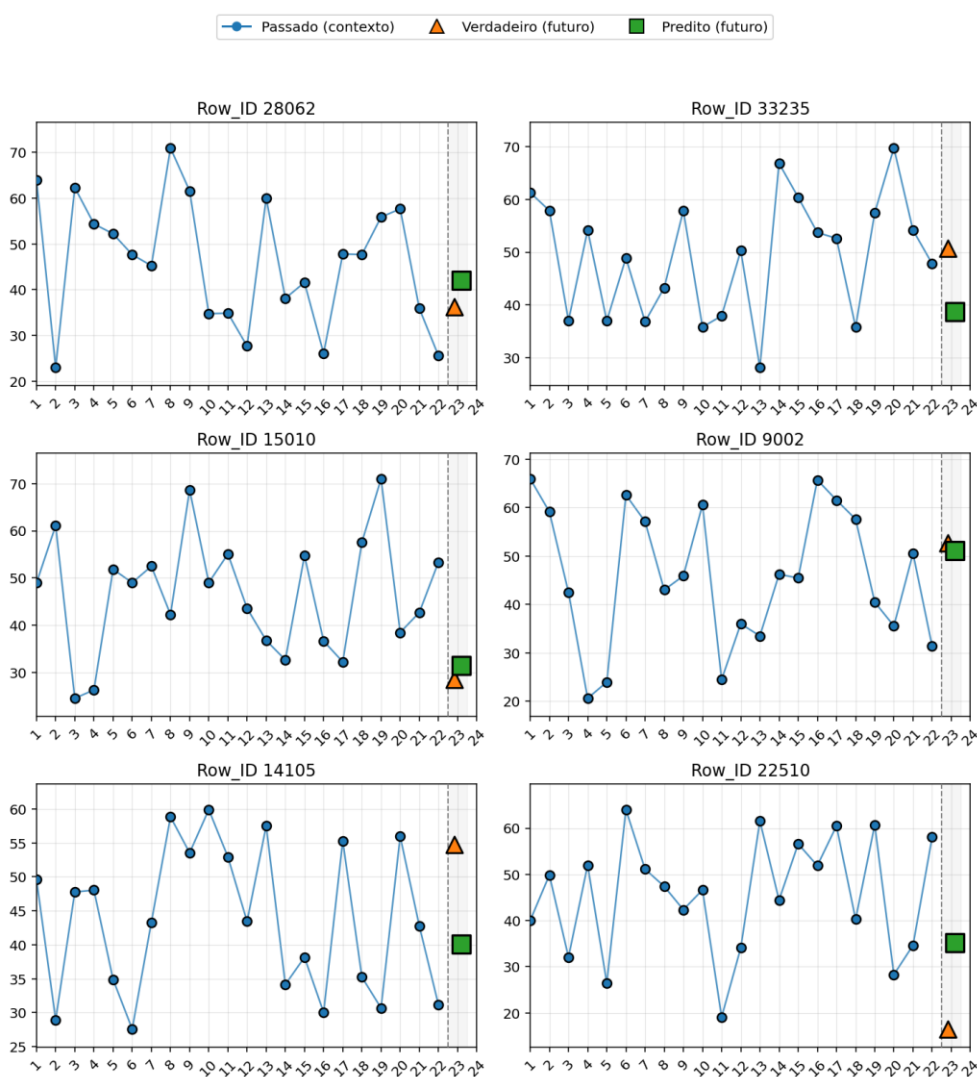
### Previsão da Rede CNN – LSTM, usando o resíduo (Delta) com apenas 1 canal

#### 1-) PREVENDO T+1 (Próximas 6 horas)

**Figura 44: Previsão: Como alvo do treinamento com resíduo delta predição T + 1 (6 horas)**

Foram utilizados 7.000 pacientes para Treinamento e Validação e 1.000 pacientes para testes

SEED = 42  
VAL\_SPLIT = 0.10  
PAST\_H = 22  
EPOCHS = 200  
FUTURE\_H = 1  
BATCH\_SIZE = 128



#### Métricas de Desempenho

=== Contagem por faixas de MAPE (TESTE) ===

< 5% : 178

5% ≤ MAPE < 10% : 169

10% ≤ MAPE < 15% : 177

15% ≤ MAPE ≤ 20% : 138

> 20% : 338

Fonte: Autor (2025)

No nono cenário da **figura 44**, observa-se os seguintes resultados:

### **Resultados no teste**

#### **Distribuição das faixas de MAPE (Erro Percentual Médio Absoluto)**

- **< 5%:** 178 pacientes
- **5% ≤ MAPE < 10%:** 169 pacientes
- **10% ≤ MAPE < 15%:** 177 pacientes
- **15% ≤ MAPE ≤ 20%:** 138 pacientes
- **> 20%:** 338 pacientes

A maior parte das previsões apresenta erro inferior a 20%, com um número considerável de previsões com erro abaixo de 10%, demonstrando que o modelo consegue uma boa precisão em muitas situações. No entanto, há também uma quantidade significativa de previsões com erro elevado (> 20%), indicando alguma dispersão no desempenho.

#### **Métricas globais**

- **MAE (Erro Médio Absoluto):** 9.8981 unidades
  - Em média, as previsões diferirem cerca de 9,90 unidades em relação ao valor real, indicando uma dispersão moderada.
- **MAPE (Erro Percentual Médio Absoluto):** 25.9536%
  - Em média, as previsões apresentam uma divergência de aproximadamente 26% em relação ao valor real, um erro considerado moderado, especialmente para previsões de horizonte de um passo.

#### **Considerações Técnicas**

O desempenho indica que o modelo tem uma boa capacidade de prever o próximo valor na sequência, com uma parcela significativa de previsões apresentando erro abaixo de 10%. Contudo, há uma dispersão considerável, evidenciada pelo elevado número de previsões com erro superior a 20%. Para aprimorar essa performance, estratégias como ajuste de hiperparâmetros, aumento de dados ou implementação de técnicas avançadas podem ser consideradas.

### 4.20 PROPOSTAS DOS CENÁRIOS TESTADOS – 10

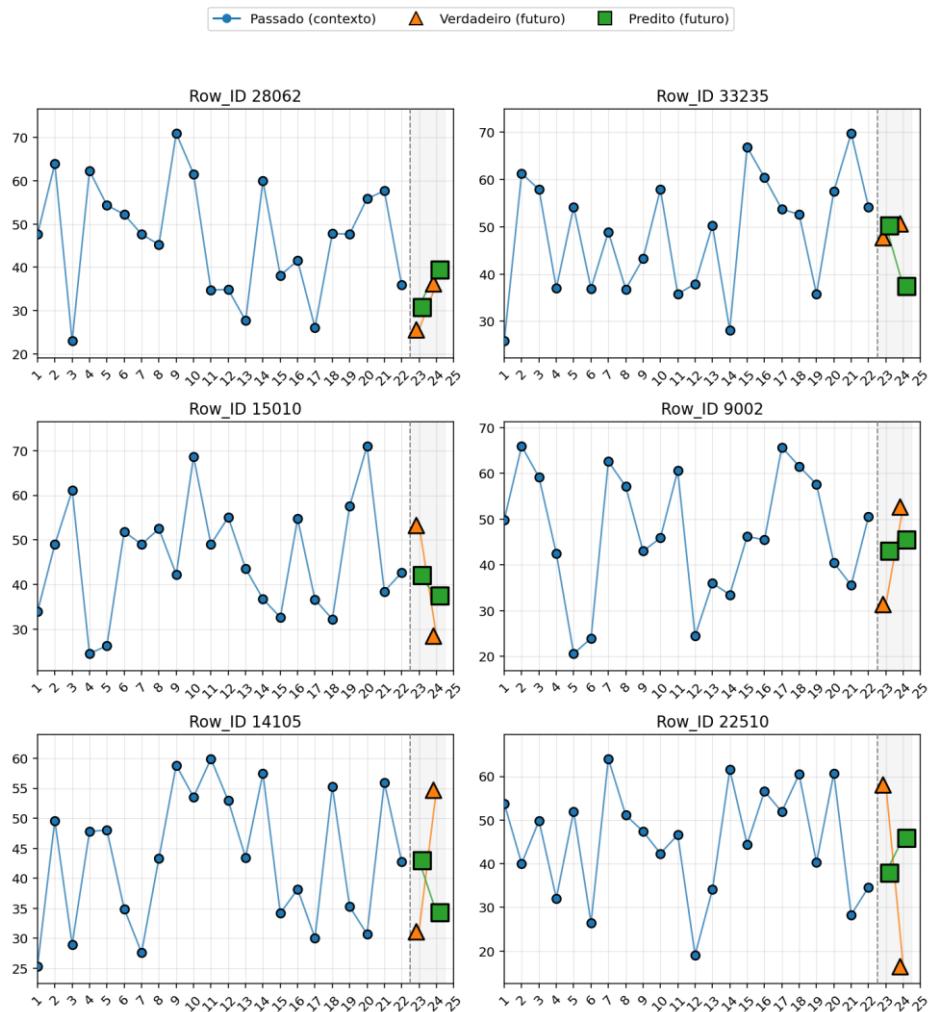
#### Previsão da Rede CNN – LSTM, usando o resíduo (Delta) com apenas 1 canal

#### 2-) PREVENDO T + 2 (Próximas 12 horas)

**Figura 45: Previsão: Como alvo do treinamento com resíduo delta predição T + 2 (12 horas)**

Foram utilizados 7.000 pacientes para Treinamento e Validação e 1.000 pacientes para testes

SEED = 42  
 VAL\_SPLIT = 0.10  
 PAST\_H = 22  
 EPOCHS = 200  
 FUTURE\_H = 2  
 BATCH\_SIZE = 128



#### Métricas de Desempenho

- < 5% : 76
- 5% ≤ MAPE < 10% : 189
- 10% ≤ MAPE < 15%: 197
- 15% ≤ MAPE ≤ 20%: 123
- > 20% : 415

Fonte: Autor (2025)

No décimo cenário da **figura 45**, observa-se os seguintes resultados:

### **Resultados no teste**

#### **Distribuição das faixas de MAPE (Erro Percentual Médio Absoluto)**

- **< 5%:** 76 pacientes
- **$5\% \leq \text{MAPE} < 10\%$ :** 189 pacientes
- **$10\% \leq \text{MAPE} < 15\%$ :** 197 pacientes
- **$15\% \leq \text{MAPE} \leq 20\%$ :** 123 pacientes
- **> 20%:** 415 pacientes

Este resultado demonstra que uma quantidade considerável de previsões (cerca de 59%) possui erro abaixo de 15%, indicando bom desempenho na maioria dos casos. Contudo, há um número expressivo de previsões com erro acima de 20%, o que sugere variabilidade na precisão, especialmente para alguns pacientes.

#### **Métricas globais**

- **MAE (Erro Médio Absoluto):** 9.8981 unidades
- **MAPE (Erro Percentual Médio Absoluto):** 25.9536%

O MAE indica que, em média, as previsões diferem aproximadamente 9,90 unidades do valor real. Já o MAPE reflete uma média de erro relativo de cerca de 26%, o que aponta para uma dispersão significativa nas previsões, especialmente nos casos mais complexos ou mais desafiadores.

#### **Considerações Técnicas**

Os resultados mostram que o modelo consegue fazer previsões relativamente precisas para uma parcela considerável de pacientes, embora haja uma dispersão considerável no desempenho geral. Para melhorar esses resultados, estratégias adicionais, como ajuste de hiperparâmetros, aumento do volume de dados ou uso de técnicas de ensemble, podem ser exploradas visando reduzir o erro e aumentar a consistência das previsões.

#### 4.21 PROPOSTA DE TESTE DO MELHOR MODELO TREINADO COM OS DADOS DO MIMIC-III USANDO OS DADOS DO DATASET MIMIC-IV

##### Previsão da Rede CNN – LSTM, usando a previsão T+1 (próximas 6 horas)

Como mencionado, o MIMIC-III consolida dados de cerca de 40 mil pacientes internados em UTIs entre 2001 e 2012 em uma estrutura de tabelas únicas, enquanto o MIMIC-IV expande significativamente esse escopo para mais de 190 mil pacientes (período de 2008 a 2019), adotando um modelo relacional mais moderno e modular que abrange não apenas a terapia intensiva, mas também dados detalhados de departamentos de emergência e internações hospitalares gerais, permitindo uma análise mais longitudinal e abrangente da jornada do paciente.

Deste modo, os resultados apresentados a seguir refletem o teste realizado com os dados do MIMIC-IV, uma vez que o treinamento do modelo de aprendizagem profunda CNN-LSTM ocorreu com o uso dos dados da base MIMIC-III. Logo, a proposta aqui foi testar o modelo com dados de pacientes de outra base de dados, não empregada na fase de treinamento, mas que foram elegíveis para calcular o IPED conforme a proposta da presente pesquisa para pacientes internados em UTI.

Para os casos testados, o modelo foi treinado com dados de 8000 pacientes usando o MIMIC-III, ou seja, com os valores do IPED calculados em 24 passos de tempo para cada paciente, e o teste ocorreu com dados de 2000 pacientes usando o IPED calculado a partir do MIMIC-IV.

**A Tabela 4: Resultados dos testes usando os dados do MIMIC-IV**

PAST_H	FUTURE_H	MAE	MAPE	R <sup>2</sup>
15	1	13,16	24,22 %	0,1386
20	1	11,48	20,76%	0,3292
21	1	10,58	19,16%	0,4313
22	1	1,60	2,87%	0,9687
23	1	0,58	0,95%	0,9982
15	2	13,20	24,23%	0,1127
20	2	11,97	21,64%	0,2571
21	2	11,41	20,68%	0,3202
22	2	10,13	18,01%	0,4623

Fonte: Autor (2025)

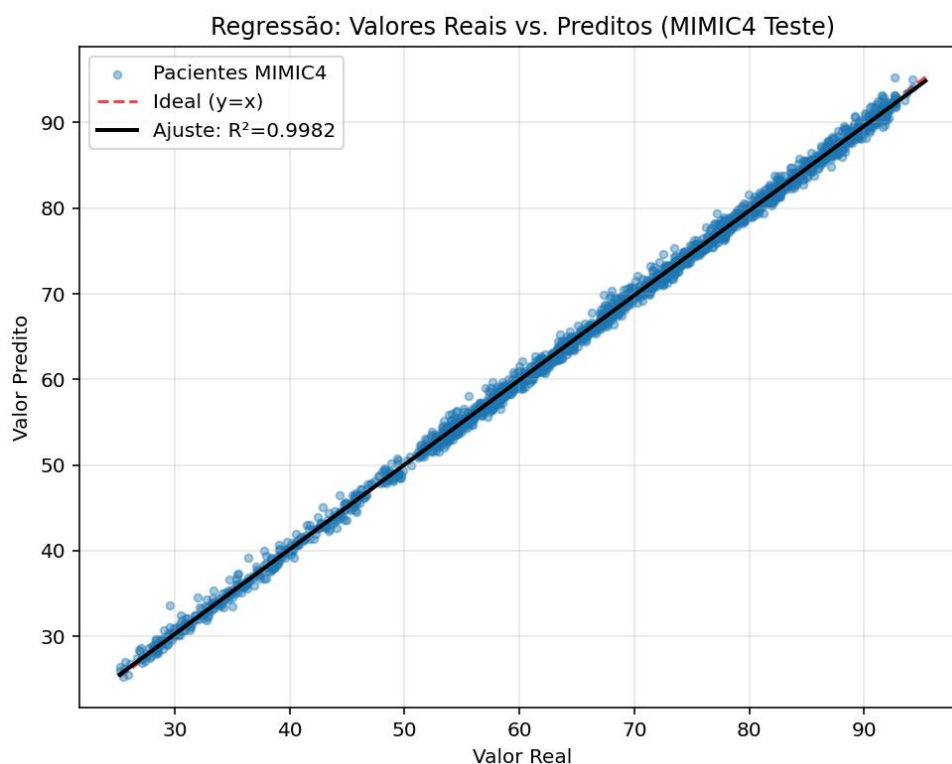
Em todos os casos, o modelo foi treinado usando 100 épocas e um valor de batch size = 128, a exemplo dos resultados obtidos e avaliados para os experimentos anteriores.

Como pode ser notado na **tabela 4**, os melhores resultados foram obtidos para a predição da deterioração no passo t+1, ou seja, para as próximas 6 horas do paciente internado. Muito

embora os resultados observados em  $t+2$  não sejam ainda muito satisfatórios, a predição nas 6 horas seguintes do paciente oferece uma oportunidade de intervenção bastante oportuna para os profissionais de saúde.

Na **figura 46** é possível notar a boa precisão obtida no melhor caso ( $t+1$ ), quando o modelo levou em conta os 23 passos passados, com base no coeficiente de regressão  $R^2$ .

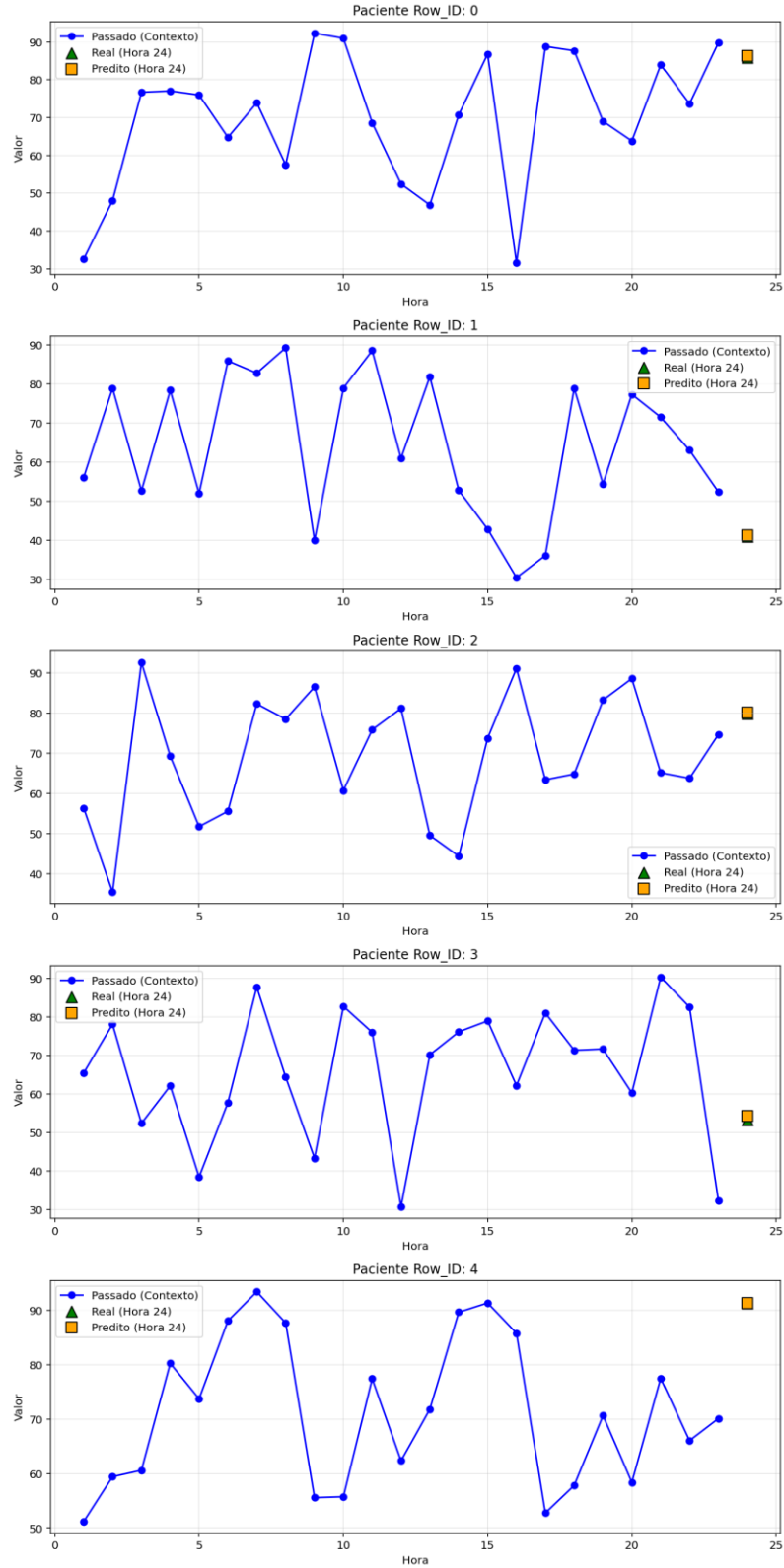
**Figura 46: Coeficiente de regressão para o melhor caso ( $t+1$ )**



Fonte: Autor (2025)

Na **figura 47** são apresentadas as curvas de predição do IPED usando dados de 5 pacientes da base MIMIC-IV. Tais resultados reforçam as métricas estatísticas alcançadas, ou seja, que a presente abordagem é capaz de contribuir de forma importante para a predição da deterioração do paciente adulto hospitalizado, tomando por base o índice ora proposto e a sua avaliação temporal com o uso de um modelo de aprendizagem profunda.

**Figura 47: Predição da deterioração do paciente (variação do IPED) para 5 pacientes da base MIMIC-IV**



Fonte: (Autor (2025))

## 4.22 CONSIDERAÇÕES FINAIS DOS RESULTADOS DOS EXPERIMENTOS COMPUTACIONAIS

O estudo observou que a aplicação de inteligência artificial (IA) híbrida na predição de deteriorações clínicas em pacientes adultos hospitalizados com o uso de registros eletrônicos de saúde (RES) apresenta diversos desafios e tendências que merecem destaque:

### **Ambiguidade terminológica e sinonímia no corpus de textos livres.**

- **Desafios:** Terminologias clínicas em saúde, muitas vezes apresentam sinônimos ou termos ambíguos, dificultando a correta interpretação dos dados se interpretados manualmente .
- **Tendência:** Potencialmente o desenvolvimento de ontologias e taxonomias padronizadas e uso de tecnologias de processamento de linguagem natural (PLN) específicas para solucionar essas questões, melhorando a consistência das informações como sugere o modelo SNOMED CT.

### **1. Necessidade de registros padronizados no corpus técnico clínico dos profissionais de saúde:**

- **Desafios:** A qualidade dos dados depende da anotação precisa de registros clínicos, que é trabalhoso e exige expertise especializada.
- **Tendência:** Automação e semi-automação da anotações e evoluções, podem além do uso de modelos de aprendizagem ativa, otimizar o tempo de dedicação do anotador, além de criar textos mais robustos para treinar modelos IA.

### **2. Ampliação e produção de modelos pré-treinados específicos para saúde para melhorar a precisão de acurácia de acertos:**

- **Desafios:** Modelos genéricos podem não captar todas as nuances dos dados clínicos, levando a baixa acurácia.
- **Tendência:** Utilização de modelos pré-treinados específicos para o domínio da saúde, como ClinicalBERT ou BioGPT, BioBERT que adaptam o entendimento às particularidades dos jargões e registros clínicos dos profissionais da saúde, aumentando a precisão das predições.

Desta forma, superar essas lacunas exige esforços multidisciplinares envolvendo especialistas em saúde, engenheiros, cientistas e analistas de dados e gestores de projetos reguladores. Investimentos em corpora de alta qualidade, tecnologias de integração de dados e

na transparência dos modelos são essenciais para maximizar os benefícios da IA na predição de deteriorações clínicas, promovendo melhorias na segurança e nos desfechos dos pacientes. (MULLAINATHAN et al., 2017 e LOURENÇO et al., 2023)

#### 4.23 PERSPECTIVAS PARA O FUTURO PARA OS AMBIENTES HOSPITALARES

- **Desenvolvimento de sistemas que integram análise textual e sinais vitais, exames laboratoriais, para suporte à decisão clínica:** Sistemas híbridos que combinam análise de texto clínico com sinais vitais em tempo real poderão fornecer uma avaliação holística do paciente. Esses sistemas possibilitarão a detecção simultânea de mudanças de condições clínicas por meio de múltiplas fontes de dados, promovendo intervenções rápidas e personalizadas na Unidade de Terapia Intensiva. A integração de dados em escala multissensorial ajudará a criar modelos preditivos mais precisos e confiáveis.
- **Ferramentas automatizadas de monitoramento contínuo e predição preditiva:** Com a evolução de sensores, IoT hospitalar e processamento de fluxo de dados, será possível criar ferramentas de monitoramento contínuo que alimentam modelos preditivos em tempo real. Essas ferramentas poderão alertar profissionais de saúde sobre riscos iminentes de deterioração com alta sensibilidade, permitindo ações preventivas antes do evento crítico. A combinação de modelos de aprendizado de máquina explicáveis e sistemas de alerta em tempo real abrirá caminho para a medicina preditiva proativa, reduzindo eventos adversos e melhorando a gestão de recursos clínicos
- O futuro reserva o desenvolvimento de **sistemas integrados, inteligentes e sensorizados** capazes de transformar a gestão clínica, especialmente em ambientes hospitalares críticos como UTIs. Investimentos em pesquisa, validação clínica e infraestrutura tecnológica serão essenciais para transformar essas inovações em práticas padrão, garantindo maior segurança e melhores desfechos para os pacientes.

Nos experimentos realizados com abordagens híbridas que combinam redes CNN e LSTM, observou-se um desempenho promissor na captura de padrões temporais complexos, especialmente em tarefas de previsão de curto a médio prazo. As configurações fornecem uma maior capacidade de modelar dependências tanto locais (via CNN) quanto de longo prazo (via LSTM), resultando em melhorias nos índices de erro em relação aos modelos isolados. Apesar disso, a dispersão dos resultados ainda evidencia dificuldades ao lidar com dados altamente variáveis e com maior horizonte de previsão, onde uma parcela significativa das previsões

apresenta erro superior a 20%. Os índices globais de erro, como MAE e MAPE, indicam que o desempenho é satisfatório em vários cenários, com potencial para aprimoramentos adicionais por meio de ajustes nos hiperparâmetros, aumento na quantidade de dados e uso de técnicas avançadas de otimização. Em suma, os modelos híbridos mostram-se uma abordagem robusta, integrando o melhor das CNNs e LSTMs, oferecendo resultados superiores às técnicas unicamente convencionais, mas ainda requerem refinamentos para atender às demandas de maior precisão em aplicações práticas mais onerosas.

#### **4.24 PERSPECTIVAS DA PROPOSTA DE PESQUISA COM A INTEROPERABILIDADE DE SISTEMAS COM REGISTROS ELETRÔNICOS DE SAÚDE.**

Pesquisas desta natureza, sugerem reflexões também sobre a interoperabilidade XAI em Registros Eletrônicos de Saúde, com sugere o estudo de **CATERSON J et al, (2024)** que refere-se à capacidade de sistemas de saúde não apenas trocarem dados brutos (como exames e sinais vitais), mas também trocarem a **explicabilidade** gerada por modelos de Inteligência Artificial. (**YUN SHEN et al, 2025**)

Tradicionalmente, a interoperabilidade foca em padrões como o **HL7 FHIR** para garantir que o sistema A entenda o dado do sistema B. No contexto da XAI, o desafio é garantir que a justificativa de uma decisão clínica tomada por uma IA no Hospital X seja interpretável e auditável pelo profissional de saúde no Hospital Y, mantendo a semântica e a confiança. (**FERNANDES PRABLU et al, 2025** e **BOSENKO I et al, 2024**)

##### **A relação Registros Eletrônicos em Saúde (RES) com a interoperabilidade**

**BALCH et al (2023)** sugerem que os Registros Eletrônicos em Saúde são a fonte primária de dados para os modelos de IA. A integração da XAI com os RES visa resolver o problema da "caixa-preta".

- **Padronização de Explicações:** Para que a XAI seja interoperável, as explicações (sejam elas visuais como mapas de calor, ou textuais como atribuição de pesos a variáveis) devem seguir padrões que possam ser incorporados aos campos do prontuário eletrônico.
- **Apoio à Decisão Clínica (CDS):** A interoperabilidade garante que os algoritmos de apoio à decisão possam operar em diferentes plataformas de prontuário, entregando

insights transparentes ao médico no ponto de cuidado (Point-of-Care).

- **Auditabilidade e Ética:** Em casos de erro médico envolvendo IA, a interoperabilidade XAI permite que a "trilha de raciocínio" da máquina esteja disponível no histórico clínico do paciente para revisão regulatória.

Nesse contexto, a **interoperabilidade** não é somente “trocar arquivos”: ela envolve camadas (técnica/sintática, semântica e organizacional). Um exemplo atual é o movimento de migração e transformação entre formatos legados (como HL7 CDA) e FHIR, com ferramentas e abordagens que visam **interoperabilidade semântica** por meio de mapeamentos verificáveis e reutilizáveis. Essas abordagens são particularmente relevantes quando você quer usar EHR para *uso primário* (cuidado) e *uso secundário* (pesquisa, gestão, qualidade), reduzindo a necessidade de migrações físicas e facilitando unificação em tempo quase real.

Há, portanto, um “acoplamento” natural entre os dois temas: **interoperabilidade cria o caminho e o vocabulário para integrar IA no fluxo clínico**, enquanto **XAI cria as evidências interpretáveis para justificar o uso da IA e apoiar decisão**. Trabalhos recentes sobre recomendações para um futuro responsável de *Toward a responsible future: recommendations for AI-enabled clinical decision support*, de LABKOFF S et al, (2024) reforçam exigências como transparência, documentação, validação, monitoramento contínuo e treinamento dos usuários — elementos que, na prática, dependem de integração técnica e de dados padronizados.

Do ponto de vista arquitetural, uma linha aplicada (muito pertinente a esta pesquisa de doutorado) onde é organizar *pipelines* preditivos em torno de padrões interoperáveis — por exemplo: EHR → normalização/terminologias → recursos FHIR → serviço de inferência → retorno ao prontuário/fluxo (CDS). A literatura recente sobre EHR e otimização de sistemas também aponta que o futuro do EHR precisa ser **adaptável e capaz de implementar IA “de forma fluida” em múltiplos contextos clínicos**, evitando que a tecnologia atrapalhe o trabalho em vez de apoiar.

Finalmente, quando o foco são modelos que usam dados tabulares longitudinais (como sinais vitais e laboratoriais), a discussão de XAI em EHR ganha corpo. Uma scoping review voltada especificamente à aplicação de XAI em pesquisas com EHR discute métodos, aplicações e limitações na tradução para a prática — o que ajuda você a justificar, na tese, por que **explicabilidade** não é “enfeite”, mas requisito de implementação segura e aceitável.

**Síntese para esta tese:** “Interoperabilidade XAI” pode ser posicionada como um eixo de **translação**: *como transformar um modelo preditivo em um serviço clínico reutilizável, auditável e aceito*, integrando padrões (p.ex., FHIR) com explicações clinicamente úteis, rastreáveis e documentadas.

#### 4.25 O CONTEXTO BRASILEIRO QUE ABORDA O TEMA DE PESQUISA

o contexto brasileiro, observa-se avanço relevante na validação de modelos preditivos de deterioração clínica com dados de prontuário eletrônico, especialmente em iniciativas multicêntricas e orientadas à implementação. Evidências incluem sistemas de alerta precoce baseados em aprendizado de máquina com validação em hospitais brasileiros e integração ao fluxo assistencial, o que destaca um diferencial prático em relação a parte da literatura internacional centrada em bases públicas como MIMIC. (**J. KOBYLARZ RIBEIRO et al. 2020**)

Entretanto, quando se considera especificamente arquiteturas híbridas profundas do tipo CNN-LSTM, a presença em publicações brasileiras de validação é mais limitada, possivelmente devido à heterogeneidade e fragmentação dos dados de EHR, à variabilidade de codificações e à irregularidade temporal dos registros, fatores que favorecem abordagens baseadas em modelos para dados tabulares que podem elevar o custo de engenharia de dados necessário para séries temporais compatíveis com redes híbridas. (**YAZDANPANA H et al, 2024**).

No Brasil, as pesquisas e iniciativas de validação com dados locais, observa-se presença de grupos nacionais que utilizam bases públicas e internacionais como a MIMIC para **benchmarking e validação metodológica** de modelos de risco em UTI, especialmente para **mortalidade**. Em particular, pesquisadores vinculados à Universidade Federal de Minas Gerais, descrito por **SANTOS et al (2018)** propuseram uma arquitetura com componentes convolucionais e recorrentes (CNN-LSTM) com **adaptação de domínio**, abordando explicitamente o desafio de **generalização** entre diferentes populações e demonstrando ganhos de desempenho ao transferir representações temporais em séries clínicas.

Há evidências de produção nacional com avaliação em MIMIC-III para mortalidade intra-hospitalar explorando heterogeneidade de dados de EHR e estratégias de combinação/ensemble, reforçando a relevância do uso de bases como MIMIC como “campo de prova” antes de validações externas em hospitais brasileiros, reduzindo o risco de viés de transferência ao migrar modelos treinados/avaliados em MIMIC para diferentes realidades

assistenciais. (GRAWE et al, 2023).

A predição precoce de deterioração clínica em pacientes adultos hospitalizados representa uma das fronteiras mais promissoras para a qualificação da assistência hospitalar no sistema público de saúde brasileiro.

A proposta de utilizar um índice clínico-laboratorial associado a um modelo híbrido de aprendizagem profunda — combinando, por exemplo, redes convolucionais (CNNs) e recorrentes (LSTMs) — permite explorar simultaneamente padrões espaciais e temporais presentes nos dados de prontuários eletrônicos de saúde. Esses modelos são capazes de aprender relações complexas entre sinais vitais, exames laboratoriais e evolução clínica ao longo do tempo, identificando trajetórias de risco que muitas vezes não são perceptíveis apenas pela avaliação humana ou por escores tradicionais.

No contexto do SUS, essa abordagem pode ser aplicada como uma ferramenta de apoio à decisão clínica integrada aos prontuários eletrônicos já utilizados em hospitais públicos. O sistema pode gerar alertas precoces de risco de deterioração, como probabilidade de instabilidade hemodinâmica, sepse, necessidade de transferência não planejada para UTI ou risco de óbito. Esses alertas, quando bem calibrados e validados localmente, podem auxiliar equipes multiprofissionais — especialmente a enfermagem, que está na linha de frente do monitoramento contínuo — a priorizar avaliações, intensificar vigilância e acionar protocolos de resposta rápida. Em um sistema público que busca eficiência e equidade, antecipar eventos adversos significa também otimizar custos e ampliar a capacidade de atendimento.

Entretanto, a implementação responsável exige atenção a fatores como qualidade dos dados, interoperabilidade entre sistemas de informação, governança de dados, ética e transparência algorítmica. A incorporação de princípios de explicabilidade (XAI) pode aumentar a confiança dos profissionais de saúde, permitindo compreender quais variáveis mais contribuem para o risco estimado e favorecendo o uso crítico e consciente das recomendações do modelo.

Assim, a pesquisa em predição de deterioração com modelos híbridos de aprendizagem profunda não deve ser vista apenas como um avanço tecnológico, mas como uma estratégia de fortalecimento do cuidado hospitalar no SUS. Ao transformar dados rotineiramente coletados em conhecimento acionável, abre-se caminho para uma assistência mais preventiva, segura e eficiente — alinhada aos princípios de integralidade, equidade e universalidade do sistema público de saúde brasileiro.

## 4.26 LIMITAÇÕES DA PESQUISA

### 1. Dados e Generalização

- **Natureza Monocêntrica:** A procedência dos dados das bases, tanto do MIMIC-III e MIMIC-IV, são do (*Medical Information Mart for Intensive Care*) do *Beth Israel Deaconess Medical Center* (BIDMC), um hospital-escola da *Harvard Medical School* localizado em *Boston*, EUA. Assim sendo, submeter o modelo a treinamento com dados de outros serviços hospitalares e outros países, são recomendados. Existe a possibilidade de variações de achados. Embora as diretrizes de construção dos algoritmos e hiperparâmetros utilizados na rede CNN-LSTM, devam ser mantidos em quaisquer outros escopos de experimentos, para que não ocorra desvio da originalidade desta pesquisa.
- **Desbalanceamento de Classes:** Eventos de deterioração clínica são, felizmente, mais raros que a estabilidade. Isso pode fazer com que o modelo seja excelente em prever estabilidade clínica dos pacientes, mas falhe em detectar os casos críticos (falsos negativos). Assim sendo, deva se considerar que o modelo de treinamento desta pesquisa, foi construído com parâmetros disponíveis na base de dados, descrita nos índices IPED e IPET nos Capítulos 3 e 4.
- **Dados Faltantes (Missing Data):** Na prática clínica, exames laboratoriais não são coletados em intervalos fixos. O modelo precisa lidar com esses "lacunas" na linha do tempo. Embora o modelo desta pesquisa, considerou na fase de limpeza, tratamento e normalização dos dados, antes da submissão aos treinamentos dos algoritmos.

### 2. Arquitetura CNN-LSTM

- **O Problema da "Caixa-Preta":** Redes profundas são complexas. Um médico pode perguntar: "Por que o índice subiu?". A dificuldade de explicar quais variáveis laboratoriais pesaram mais na decisão é uma limitação de **interpretabilidade**. Esta questão foi abordada e considerada na fase de definição dos pesos de cada índice IPED e IPET com regressão logística nos dados disponíveis na amostra disponível.
- **Dependência Temporal:** O componente LSTM assume que o passado prediz o futuro. Mas e se uma intervenção médica (como uma cirurgia de emergência) mudar o curso clínico subitamente? Esta questão não foi explorada nesta pesquisa, sugerimos que

outros estudos possam aprofundar-se.

### 3. Implementação e Aspectos Éticos

- **Fadiga de Alarmes:** Se o modelo for muito sensível, ele pode gerar alertas constantes para pacientes que não estão em risco real, levando a equipe de enfermagem e médica a ignorar o sistema. Os ajustes de parâmetros nos equipamentos disponíveis no ambiente hospitalar, são sugeridos aos novos estudos.
- **Viés Algorítmico Demográfico:** O modelo pode performar de forma diferente dependendo da idade, etnia ou comorbidades se essas subpopulações não estiverem bem representadas no treino dos dados disponíveis;
- **Viés de Validação do Modelo:** Embora este estudo trabalhou com dados reais de pacientes anonimizados, sugere-se que o processo de validação metodológica com especialistas, seja aplicado a quaisquer modelos, antes de serem implementados em sistemas de PEP (Prontuário Eletrônico do Paciente) ou em sistemas de apoio a decisão.
- **Textos livres dos registros eletrônicos dos pacientes:** O desempenho do índice IPET que observou textos livres, como anotações e evoluções dos profissionais de saúde, não demonstraram resultados satisfatórios neste estudo. O impacto de pesos deste índice de texto, estiveram abaixo de 0,80%. Portanto, este estudo sugere investigações com técnicas de NLP (*Natural Language Process*) e pacotes de dados pré treinados do tipo BERT em larga escala.

#### 4.27 CONTRIBUIÇÕES DA PESQUISA (Pontos Fortes)

Os modelos de pesquisa baseados nas bases **MIMIC-III** e **MIMIC-IV** destacam-se por democratizar o acesso a dados clínicos reais e de alta complexidade, permitindo avanços significativos em medicina de precisão com registros eletrônicos de saúde com inteligência artificial.

##### 1. Alta Granularidade e Diversidade de Dados

Diferente de bases administrativas simples, o MIMIC oferece uma visão microscópica da jornada do paciente no ambiente de práticas hospitalares;

- **Séries Temporais:** Registros de sinais vitais em espaços curtos de tempos, com

informações de desfechos de altas e óbitos.

- **Multimodalidade:** Integração de resultados laboratoriais;

## 2. Acesso Aberto e Reprodutibilidade

O modelo de "esquema de acesso permissivo (acessibilidade)" dos dados, é um dos maiores pilares da base de dados utilizada:

- **Comunidade Global:** Por ser gratuito para pesquisadores que cumprem os requisitos éticos, ele fomenta uma colaboração mundial, permitindo que diferentes equipes validem ou refutem descobertas em cima da mesma base. As bases funcionam como o "padrão ouro" para o treinamento de modelos preditivos em saúde:

- **Benchmarks de Performance:** É a base principal para testar algoritmos de predição de mortalidade, detecção precoce de sepse e estimativa de tempo de permanência hospitalar.
- **Volume de Dados:** O MIMIC III e MIMIC IV expandiu significativamente as possibilidades de pesquisa, cobrindo mais de uma década de internações e incluindo dados de condições críticas do o que reduz vieses de amostras pequenas.

## 4. Realismo Clínico

Diferente de ensaios clínicos controlados, o MIMIC reflete a **prática médica real**:

- **Complexidade:** Captura as nuances de pacientes com múltiplas comorbidades e as variações nas decisões terapêuticas tomadas sob pressão em UTIs.
- **Atualidade:** O MIMIC-IV traz dados contemporâneos, incorporando práticas médicas e sistemas de prontuário mais modernos que a versão anterior.
- **Por que a escolha do modelo híbrido utilizado?**

O modelo combina a capacidade das redes neurais CNN para extrair padrões complexos com habilidades das redes LSTM de lidar com sequências temporais.

## 5.0 CONCLUSÕES

A partir dos resultados obtidos na pesquisa, em propor um novo índice de predições de deterioração clínica laboratorial de pacientes adultos, com uso da aprendizagem de máquina para apoiar a tomada de decisão de médicos e enfermeiros e nortear condutas em um ambiente hospitalar. Foi inicialmente produzida uma importante revisão de escopo com o tema de pesquisa, explorou-se as bases de dados do MIMIC-III, com dados estruturados e dados não estruturados presentes nos registros eletrônicos de saúde.

Foi realizado a limpeza e o tratamento dos registros eletrônicos, seguindo as propostas de Descoberta de Conhecimento em Bases de Dados (KDD), criou-se os modelos de treinamento supervisionado. Demonstrou-se a arquitetura proposta das redes neurais. Construiu-se então os modelos matemáticos e suas equações para cálculo dos pesos das variáveis.

A pesquisa conseguiu atingir seus objetivos quando:

- Desenvolveu-se um novo índice para a predição futura da deterioração clínica laboratorial de pacientes adultos, a partir de sinais vitais e exames laboratoriais de cada paciente, considerando as variáveis estatisticamente mais relevantes;

- Criado um banco de dados que representa a variação temporal do índice para cada paciente;

- Desenvolvidos, treinados e validados ao menos um modelo de aprendizado de máquina, a partir do banco de dados criado, capaz de prever a deterioração clínica laboratorial de um paciente, a partir do uso do índice calculado em intervalos de tempos passados;

Testado o desempenho do modelo de aprendizagem de máquina treinado, a partir de casos de pacientes não empregados na fase de treino e validação;

- Comparado desempenho do índice com dados reais de predição de deterioração clínica, encontrados nas bases de dados além de projetar o risco de mortalidade dos pacientes a partir do modelo de predição proposto;

- Decidido pelo uso final dos experimentos com um modelo híbrido de redes neurais com CNN e LSTM, apresentados os resultados e suas métricas de desempenho.

A modelo realizou ajustes em 10 cenários diferentes, para avaliação de seus desempenhos, com ajustes de hiperparâmetros em “ *Seed, Past\_H, Past\_H, Future\_H, Val\_Split, Epochs e Batch\_size* “, onde foi possível demonstrar a predição de eventos futuros com uma série temporal retroativa de registros e treinamentos para prever a sexta (6) hora,

décima segunda (12), décima oitava (18) hora e vigésima quarta (24) hora. O desempenho obteve métricas globais de MAE = 9,36 - 9,89 e MAPE = 24,19 - 25,96%.

O modelo híbrido proposto apresentou desempenho e resultados muito próximos, quando submetido a duas bases de dados diferentes com pacientes diferentes, como no MIMIC-III e o MIMIC-IV, em condições e cenários complexos e heterogêneos. Em todos os testes, o modelo foi treinado com pelo menos 100 épocas e um valor de batch size = 128. Os melhores resultados foram obtidos para a predição da deterioração no passo  $t+1$ , ou seja, para as próximas 6 horas do paciente. Desta forma a predição nas 6 horas seguintes do paciente, oferece uma oportunidade de intervenção bastante oportuna para os profissionais de saúde. Será apresentado a comunidade hospitalar, logo que os resultados na pesquisa sejam publicados, uma análise de viabilidade de interoperação e integração do índice na rotina de sistemas de prontuários eletrônicos hospitalares. Assim, entende-se que esta pesquisa apresenta resultados robustos e fiéis por ter explorado informações reais de pacientes anonimizados.

Desta forma, estudos como este, podem impulsionar novas oportunidades, para prever tendências e desafios de pesquisa, ou que novos pesquisadores possam contestar ou agregar novas experiências dos resultados identificados até aqui.

Por fim, sugerimos que este estudo, apresenta um potencial de ineditismo explorado na condução deste trabalho, que projeta contribuições científicas impactantes para a comunidade acadêmica, e que seus resultados podem antecipar eventos e incidentes clínicos no ambiente hospitalar para a segurança e proteção dos pacientes e das instituições.

## 6.0 ANEXOS

## Anexo01 – Print da página principal da publicação


Proposed use of data mining techniques in healthcare for knowledge generation in patient care: an integrative review



Oliveira, R. R. de, Santos, S. L. B. dos, & Sassi, R. J. (2023, August 14). Proposed use of data mining techniques in healthcare for knowledge generation in patient care: an integrative review. *II SEVEN INTERNATIONAL CONGRESS OF HEALTH*. <https://doi.org/10.56238/homeIIsevenhealth-054>

## Anexo02 – Print da página principal da publicação

Uso de aprendizado de máquina em registros eletrônicos de Saúde para prever deteriorações clínicas e mortalidade em pacientes adultos hospitalizados



REVISTA DE GESTÃO E  
SECRETARIADO

ISSN: 2178-9010  
DOI: <http://doi.org/10.7769/gesec.v16i4.4882>

**Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados**

**Using Machine Learning in Electronic Health Records to Predict Clinical Deterioration and Mortality in Hospitalized Adult Patients**

**Uso del Aprendizaje Automático en Registros Médicos Electrónicos para Predecir el Deterioro Clínico y la Mortalidad en Pacientes Adultos Hospitalizados**


Reinaldo Ribeiro de Oliveira<sup>1</sup>  
Cleber Gustavo Dias<sup>2</sup>

**Resumo**

**Introdução:** O processo de hospitalização de pacientes envolve cuidados especializados, com equipes multiprofissionais de saúde, focados na recuperação e retomada do paciente ao estado de equilíbrio. Na hospitalização, pode acontecer deteriorações clínicas. Os Registros Eletrônicos de Saúde (RES) são, em sua forma mais simples, versões digitais dos prontuários dos pacientes. Eles disponibilizam informações instantaneamente, "quando e onde forem necessárias" **Justificativa:** Entendemos oportunidades de propor e produzir um experimento tecnológico computacional para condicionar um algoritmo matemático, aprender e prever tendências de complicações de pacientes hospitalizados. **Objetivo:** Propor com aprendizado de máquina em registros eletrônicos de saúde, algoritmos de treinamento para prever deteriorações clínicas em pacientes adultos hospitalizados. **Metodologia:** A Fase1 primeira foi produzida uma revisão de escopo, na Fase2: foram construídos os experimentos com elaboração dos algoritmos de Machine Learning, onde submetemos a base de dados MIMICIV com 223.452 pacientes com 20.000 registros tratados de pacientes para a pesquisa. **Resultados:** Foram produzidos 6 experimentos computacionais com treinamentos supervisionados, usando técnicas de Regressão Logística, Random Forest e XGboost, cujos resultados demonstram

<sup>1</sup>Mestre em Enfermagem, Centro Universitário São Camilo, São Paulo, São Paulo, Brasil.  
E-mail: [reinaldo.gestao.oliveira@gmail.com](mailto:reinaldo.gestao.oliveira@gmail.com) Orcid: <https://orcid.org/0000-0002-1219-0652>

<sup>2</sup>Pós-doutor em Engenharia Elétrica, Universidade de São Paulo, São Paulo, São Paulo, Brasil.  
E-mail: [diascg2019@gmail.com](mailto:diascg2019@gmail.com) Orcid: <https://orcid.org/0000-0002-4232-2409>



Revista de Gestão e Secretariado – GeSec, V. 16, N. 4, P. 01-23, 2025  
São José dos Pinhais, Paraná, Brasil.

1

Oliveira, R. R. de, & Dias, C. G. (2025). Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados. *Revista de Gestão e Secretariado*, 16(4), e4882. <https://doi.org/10.7769/gesec.v16i4.4882>

### Anexo03 – Softwares, Bibliotecas e Programas Aplicativos usados nos experimentos computacionais

Software	Descrição	Utilização	URL
Jupyter Notebook	IDE de desenvolvimento em Python	Desenvolvimento de códigos Python	<a href="https://jupyter.org/">https://jupyter.org/</a>
Colab Pro	IDE de desenvolvimento em Python	Desenvolvimento de códigos Python	<a href="https://colab.google/">https://colab.google/</a>
NumPy	Biblioteca para operações numéricas e manipulação de arrays.	Tratamento de dados numéricos, cálculos matemáticos.	<a href="https://numpy.org/">https://numpy.org/</a>
Pandas	Biblioteca para manipulação e análise de dados estruturados.	Leitura, escrita, limpeza e manipulação de datasets.	<a href="https://pandas.pydata.org/">https://pandas.pydata.org/</a>
Dask	Biblioteca de computação paralela para processamento de grandes volumes de dados.	Processamento de big data além da memória RAM.	<a href="https://dask.org/">https://dask.org/</a>
Apache Spark (via PySpark)	Plataforma de processamento de big data distribuída, acessível com Python.	Análises em larga escala, processamento distribuído.	<a href="https://spark.apache.org/">https://spark.apache.org/</a>
Matplotlib	Biblioteca para criação de gráficos estatísticos e científicos.	Visualização de dados, gráficos de linhas, barras, etc.	<a href="https://matplotlib.org/">https://matplotlib.org/</a>
Seaborn	Biblioteca baseada em Matplotlib, para visualizações estatísticas sofisticadas.	Gráficos avançados com visual atraente.	<a href="https://seaborn.p">https://seaborn.p</a>
Plotly	Biblioteca para gráficos interativos e de alta qualidade.	Visualizações dinâmicas e interativas.	<a href="https://plotly">https://plotly</a>
Bokeh	Biblioteca para visualizações interativas em navegadores.	Criação de dashboards e gráficos interativos.	<a href="https://bokeh.org/">https://bokeh.org/</a>
Scikit-learn	Biblioteca de machine learning para análise preditiva.	Modelagem, classificação, regressão, clustering.	<a href="https://scikit-learn.org/stable/">https://scikit-learn.org/stable/</a>
Statsmodels	Biblioteca para análise estatística e modelagem de dados.	Modelos estatísticos, testes e exploração.	<a href="https://www.statsmodels.org/s">https://www.statsmodels.org/s</a> <a href="https://www.statsmodels.org/s">table/</a>
Gensim	Biblioteca para modelagem de tópicos, Word2Vec, e NLP baseada em tópicos.	Modelagem de tópicos, vetores de palavras, similaridade.	<a href="https://radimrehurek.com/gen">https://radimrehurek.com/gen</a> <a href="https://radimrehurek.com/gen">sim/</a>

spaCy	Biblioteca avançada para NLP, com foco em desempenho e produção.	Reconhecimento de entidades, análise sintática, vetores de palavras.	<a href="https://spacy.io/">https://spacy.io/</a>
NLTK	Biblioteca para tarefas de NLP, incluindo tokenização, análise sintática, etc.	Processamento de linguagem natural básica.	<a href="https://www.nltk.org/">https://www.nltk.org/</a>
Altair	Biblioteca para visualizações declarativas com códigos concisos.	Visualizações estatísticas e exploratórias.	<a href="https://altair-viz.github.io/">https://altair-viz.github.io/</a>
Keras	API de alto nível para redes neurais, que roda em TensorFlow, Theano ou CNTK.	Construção e treinamento de redes neurais profundas.	<a href="https://keras.io/">https://keras.io/</a>
TensorFlow	Biblioteca de código aberto para aprendizado de máquina, especialmente deep learning.	Desenvolvimento de modelos de deep learning.	<a href="https://tensorflow.org/">https://tensorflow.org/</a>
PyTorch	Biblioteca de aprendizado profundo de código aberto, com ênfase em pesquisa e desenvolvimento.	Construção de modelos de deep learning e treinamento dinâmico.	<a href="https://pytorch.org/">https://pytorch.org/</a>
Transformers	Biblioteca para modelos de processamento de linguagem natural baseados em Transformers.	Modelos pré-treinados como BERT, GPT, RoBERTa, etc.	<a href="https://huggingface.co/transformers/">https://huggingface.co/transformers/</a>
LightGBM	Biblioteca de aprendizado de máquina baseada em gradient boosting, otimizada para velocidade e eficiência.	Problemas de classificação e regressão.	<a href="https://lightgbm.readthedocs.io/">https://lightgbm.readthedocs.io/</a>
XGBoost	Biblioteca de gradient boosting otimizada, muito usada em competições.	Problemas de classificação, regressão, ranking.	<a href="https://xgboost.readthedocs.io/">https://xgboost.readthedocs.io/</a>
CatBoost	Biblioteca de boosting de gradiente desenvolvida pela Yandex, fácil de usar com dados categóricos.	Classificação, regressão, ranking.	<a href="https://catboost.ai/">https://catboost.ai/</a>
UMAP	Técnica de redução de dimensionalidade para visualização de grandes dados.	Visualização de dados de alta dimensão, agrupamentos.	<a href="https://umap-learn.readthedocs.io/">https://umap-learn.readthedocs.io/</a>
t-SNE	Técnica de redução de dimensionalidade para	Visualização de clusters, exploração de	Disponível em Scikit-learn.

	visualização de dados de alta dimensão.	dados.	
Spark MLlib	Biblioteca de machine learning para Apache Spark.	Modelagem de grandes dados distribuídos.	<a href="https://spark.apache.org/mllib/">https://spark.apache.org/mllib/</a>
MapReduce (via Hadoop)	Modelo de processamento paralelo para processamento massivo de dados.	Processamento de big data em cluster.	<a href="https://hadoop.apache.org/">https://hadoop.apache.org/</a>
Ms Excel	Planilha Eletrônica	Desenvolvedor Gráficos simples e analisar os resultados dos experimentos.	<a href="https://www.microsoft.com/pt-br/Microsoft-365/excel">https://www.microsoft.com/pt-br/Microsoft-365/excel</a>
SQL (via SQLAlchemy)	Ferramenta para consulta e manipulação de bancos de dados relacionais.	Extração, transformação e carga (ETL) de dados estruturados.	<a href="https://www.sqlalchemy.org/">https://www.sqlalchemy.org/</a>

Fonte: Autor (2025)

Conforme a tabela acima, temos o *NumPy*, uma biblioteca fundamental para operações numéricas em *Python*, permitindo manipulação eficiente de *Arrays* e matrizes, além de oferecer funções matemáticas avançadas necessárias em cálculos científicos. *Pandas* é essencial para a manipulação e análise de dados estruturados, facilitando a leitura, escrita, limpeza e transformação de grandes conjuntos de dados em formato tabular. *Dask* amplia esse conceito ao possibilitar o processamento paralelo e distribuído de volumes de dados que ultrapassam a memória RAM do sistema, enquanto o *PySpark* fornece uma interface *Python* para o *Apache Spark*, permitindo o processamento de big data em ambientes distribuídos. Para tarefas de aprendizado de máquina, *Scikit-learn* oferece uma vasta gama de algoritmos para classificação, regressão, *clustering* e validação de modelos, sendo uma ferramenta padrão na área. *XGBoost*, *LightGBM* e *CatBoost* são exemplos de algoritmos de *boosting*, altamente eficientes para problemas preditivos em grandes volumes de dados, com desempenho otimizado e suporte a variáveis categóricas. *TensorFlow* e *PyTorch* são bibliotecas de deep learning que suportam a construção de modelos complexos, com *TensorFlow* focado em produção e escalabilidade, e *PyTorch* oferecendo maior flexibilidade para pesquisa e desenvolvimento dinâmico. *Keras* funciona como uma interface de alto nível para redes neurais, facilitando a criação rápida de modelos sobre *TensorFlow* ou *Theano*.

## 7.0 REFERÊNCIAS

- ABBAS, Q., JEONG, W., & LEE, S. W. (2025). Explainable AI in Clinical Decision Support Systems: A Meta-Analysis of Methods, Applications, and Usability Challenges. *Healthcare*, 13(17), 2154. <https://doi.org/10.3390/healthcare13172154>
- ABDEL, 'T. I. et al. "A new model for tracking and detection of deterioration of vital signs based on artificial neural network." (2019).
- AHMAD, P. N., SHAH, A. M., & LEE, K. Y. A Review on Electronic Health Record Text-Mining for Biomedical Name Entity Recognition in Healthcare Domain. In *Healthcare (Switzerland)*, v. 11, n. 9, p. 1268, 2023. Disponível em: < <https://doi.org/10.3390/healthcare11091268> >. Acesso em: 28/10/2025.
- AIKODON, N., ORTEGA-MARTORELL, S., & OLIER, I. (2024). Predicting decompensation risk in intensive care unit patients using machine learning. *Algorithms*, 17(1), 6. <https://doi.org/10.3390/a17010006> Acesso em 16/01/2026.
- AKEL, M. A., CAREY, K. A., WINSLOW, C. J., CHURPEK, M. M., & EDELSON, D. P. Less is more: Detecting clinical deterioration in the hospital with machine learning using only age, heart rate, and respiratory rate. *Resuscitation*, v. 168, p. 6–10, 2021. Disponível em: < <https://doi.org/10.1016/j.resuscitation.2021.08.024> >. Acesso em: 28/10/2025.
- ALAM, F., AFZAL, M., & MALIK, K. M. Comparative Analysis of Semantic Similarity Techniques for Medical Text. *International Conference on Information Networking (ICOIN)*, Barcelona, p. 106–109, 2020. Disponível em: < <https://doi.org/10.1109/ICOIN48656.2020.9016574> >. Acesso em: 28/10/2025.
- ALSHWAHEEN, T. I., HAU, Y. W., ASS'AD, N., & ABUALSAMEN, M. M. A Novel and Reliable Framework of Patient Deterioration Prediction in Intensive Care Unit Based on Long Short-Term Memory-Recurrent Neural Network. *IEEE Access*, v. 9, p. 3894–3918, 2021. DOI: 10.1109/ACCESS.2020.3047186. Disponível em: < <https://doi.org/10.1109/ACCESS.2020.3047186> >. Acesso em: 28/10/2025.
- ALSHWAHEEN, T. I., MOGHBEL, M., HAU, Y. W., & OOI, C. Y. Use of learning approaches to predict clinical deterioration in patients based on various variables: a review of the literature. *Artif Intell*, v. 55, n. 2, p. 1055–1084, 2022. Disponível em: < <https://doi.org/10.1007/s10462-021-09982-2> >. Aceswso em: 28/10/2025.
- ALTMAN, D. G., BOOTH, A., WHITLOCK, E. (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Espanola de Nutricion Humana y Dietetica*, 20(2), 148–160. <https://doi.org/10.1186/2046-4053-4-1> Acesso em 16/01/2026
- ALVES, T., LAENDER, A., VELOSO, A., & ZIVIANI, N. (2018). Dynamic Prediction of ICU Mortality Risk Using Domain Adaptation. 2018 *IEEE International Conference on Big Data (Big Data)*, 1328–1336. <https://doi.org/10.1109/BigData.2018.8621927> Acesso em 16/01/2026 Acesso em 02/02/2026.
- ANTUNES, J. L. F., & CARDOSO, M. R. A. (2015). Uso da análise de séries temporais em estudos epidemiológicos. *Epidemiologia e Serviços de Saúde*, 24(3), 565–576. <https://doi.org/10.5123/S1679-49742015000300024> Acesso em 02/02/2026.
- ARMAIGNAC, D. L., et al (2025). Guiding Principles for Data Sharing and Harmonization: Results of a Systematic Review and Modified Delphi From the Society of Critical Care Medicine Data Science Campaign. *Critical Care Medicine*, v.53, n. 3, p. e619–e631, 2025. DOI: 10.1097/CCM.0000000000006578 <https://doi.org/10.1097/CCM.0000000000006578> >. Acesso em: 28/10/2025.
- ASHIMWE SB, VITTINGHOFF E, WHOOLEY M. (2020) Vital Signs Data and Probability of Hospitalization, Transfer to Another Facility, or Emergency Department Death Among Adults Presenting for Medical Illnesses to the Emergency Department at a Large Urban Hospital in the United States. *J Emerg Med*. 2020 Apr;58(4):570-580. <https://doi.org/10.1016/j.jemermed.2019.11.020> Acesso em 02/02/2026.

AUNG, Y. Y. M., WONG, D. C. S., & TING, D. S. W. (2021). The promise of artificial intelligence: A review of the opportunities and challenges of artificial intelligence in healthcare. In *British Medical Bulletin* (Vol. 139, Issue 1, pp. 4–15). Oxford University Press. <https://doi.org/10.1093/bmb/ldab016> Acesso em 16/01/2026.

BAEK, S., JANG, J., & YOON, S. (2019). End-to-end blood pressure prediction via fully convolutional networks (referência frequente em pipelines com MIMIC ). *IEEE Access*, 7, 185458–185468. <https://doi.org/10.1109/ACCESS.2019.2960844> Acesso em 16/01/2026.

BAKER, S., XIANG, W., & ATKINSON, I. (2020). Continuous and automatic mortality risk prediction using vital signs in the intensive care unit: a hybrid neural network approach. *Scientific Reports*, 10(1). <https://doi.org/10.1038/s41598-020-78184-7> Acesso em 16/01/2026

BAKER, S., XIANG, W., & ATKINSON, I. (2020). Continuous and automatic mortality risk prediction using vital signs in the intensive care unit: A hybrid neural network approach. *Scientific Reports*, 10, 1–12. <https://doi.org/10.1038/s41598-020-78184-7> Acesso em 16/01/2026.

BAKER, S., et al. (2022). A computationally efficient CNN-LSTM neural network for estimation of blood pressure from features of electrocardiogram and photoplethysmogram waveforms. *Knowledge-Based Systems*, 109151. <https://doi.org/10.1016/j.knosys.2022.109151> Acesso em 16/01/2026.

BALCH, J. A., RUPPERT, M. M., LOFTUS, T. J., GUAN, Z., REN, Y., UPCHURCH, G. R., OZRAZGAT-BASLANTI, T., RASHIDI, P., & BIHORAC, A. (2023). Machine Learning-Enabled Clinical Information Systems Using Fast Healthcare Interoperability Resources Data Standards: Scoping Review. *JMIR Medical Informatics*, 11, e48297–e48297. <https://doi.org/10.2196/48297> Acesso em 16/01/2026.

BATTULA, H., LIU, J., SRIVASTAVA, J. Enhancing In-Hospital Mortality Prediction Using Multi-Representational Learning with LLM-Generated Expert Summaries. 2024. <https://doi.org/10.48550/arXiv.2411.16818> Acesso em 16/01/2026.

BELL, D., BAKER, J., WILLIAMS, C., & BASSIN, L. (2021). A Trend-Based Early Warning Score Can Be Implemented in a Hospital Electronic Medical Record to Effectively Predict Inpatient Deterioration. *Critical Care Medicine*, 49(10), E961–E967. <https://doi.org/10.1097/CCM.0000000000005064> Acesso em 16/01/2026.

BERKOWITZ, J. S., SRINIVASAN, A., MIGUEL ACITORES CORTINA, J., FATAPOUR, Y., & TATONETTI, N. P. (2024). Biomedical Text Normalization through Generative Modeling. <https://doi.org/10.1101/2024.09.30.24314663> Acesso em 16/01/2026.

BLYTHE R, PARSONS R, WHITE NM, COOK D, MCPHAIL S. A scoping review of real-time automated clinical deterioration alerts and evidence of impacts on hospitalised patient outcomes. *BMJ Qual Saf*. 2022 Oct;31(10):725-734. Epub 2022 Jun 22. DOI: [10.1136/bmjqs-2021-014527](https://doi.org/10.1136/bmjqs-2021-014527) Acesso em 16/01/2026.

BOSE, P., SRINIVASAN, S., SLEEMAN, W. C., PALTA, J., KAPOOR, R., & GHOSH, P. (2021). A survey on recent named entity recognition and relationship extraction techniques on clinical texts. In *Applied Sciences* (Switzerland) (Vol. 11, Issue 18). MDPI. <https://doi.org/10.3390/app11188319> Acesso em 16/01/2026.

BOSENKO, I., RANDMAA, R., PIHO, G., & ROSS, P. (2024). Interoperability of health data using FHIR Mapping Language: transforming HL7 CDA to FHIR with reusable visual components. *Frontiers in Digital Health*, 6. <https://doi.org/10.3389/fdgth.2024.1480600> Acesso em 16/01/2026.

BOULITSAKIS LOGOTHETIS, S., GREEN, D., HOLLAND, M., & AL MOUBAYED, N. (2023). Predicting acute clinical deterioration with interpretable machine learning to support emergency care decision making. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-40661-0> Acesso em 16/01/2026.

BREKKE, I. J., PUNTERVOLL, L. H., PEDERSEN, P. B., KELLETT, J., & BRABRAND, M. (2019). The value of vital sign trends in predicting and monitoring clinical deterioration: A systematic review. *PLOS ONE*, 14(1), e0210875. <https://doi.org/10.1371/journal.pone.0210875> Acesso em 16/01/2026.

CARRASCO-RIBELLES, L. A., LLANES-JURADO, J., GALLEGO-MOLL, C., CABRERA-BEAN, M., MONTEAGUDO-ZARAGOZA, M., VIOLÁN, C., & ZABALETA-DEL-OLMO, E. (2023). Prediction models using artificial intelligence and longitudinal data from electronic health records: a systematic methodological review. In *Journal of the American Medical Informatics Association* (Vol. 30, Issue 12, pp. 2072–2082). Oxford University Press. <https://doi.org/10.1093/jamia/ocad168> Acesso em 16/01/2026.

CARUANA, A., BANDARA, M., MUSIAL, K., CATCHPOOLE, D., & KENNEDY, P. J. (2023). Machine learning for administrative health records: A systematic review of techniques and applications. *Artificial Intelligence in Medicine*, 144, 102642. <https://doi.org/10.1016/j.artmed.2023.102642> Acesso em 16/01/2026.

CATERSON, J., LEWIN, A., & WILLIAMSON, E. (2024). The application of explainable artificial intelligence (XAI) in electronic health record research: A scoping review. *DIGITAL HEALTH*, 10. <https://doi.org/10.1177/20552076241272657> Acesso em 16/01/2026.

CHANG, E., MOSTAFA, J., YU, F., PFAFF, E. R., POTENZIANI, D. D., & ROBASKY, K. (2022). Cohort identification from free-text clinical notes using snomed ct's semantic relations. Acesso em 16/01/2026.

CHANG, H.M., HUANG, E.-W., HOU, • I-CHING, LIU, H.-Y., FANG-, •, LI, S., & CHIOU, S.-F. (2022). Using a Text Mining Approach to Explore the Recording Quality of a Nursing Record System

CHEN J, QI TD, VU J, WEN Y. A deep learning approach for inpatient length of stay and mortality prediction. *J Biomed Inform.* 2023 Nov;147:104526. Epub 2023 Oct 17. <https://doi.org/10.1016/j.jbi.2023.104526> Acesso em 16/01/2026.

CHEN, F., WANG, L., HONG, J., JIANG, J., & ZHOU, L. (2024). Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. *Journal of the American Medical Informatics Association*, 31(5), 1172–1183. <https://doi.org/10.1093/jamia/ocae060> Acesso em 16/01/2026.

CHEN, K., ZHOU, X., BAO, Z., SKIBNIEWSKI, M. J., & FANG, W. (2024). Artificial intelligence in infrastructure construction: A critical review. In *Frontiers of Engineering Management*. Higher Education Press Limited Company. <https://doi.org/10.1007/s42524-024-3128-5> Acesso em 16/01/2026.

Chen, X., et al. (2024). Acute complication prediction and diagnosis model based on MIMIC-IV (CLSTM-BPR). *Tsinghua Science and Technology*. <https://doi.org/10.26599/TST.2023.9010103> Acesso em 16/01/2026.

CHENG, A. C., DUDA, S. N., MCGUFFIN, K., MCEVER, M., TAYLOR, R., REZNICZEK, G. A., MARTIN, A., MORALES, E., & HARRIS, P. A. (2025). Supporting rapid innovation in research data capture and management: the REDCap external module framework. *Journal of the American Medical Informatics Association*. <https://doi.org/10.1093/jamia/ocaf073> Acesso em 16/01/2026.

CHIU, C.-C., Wu, C.-M., Chien, T.-N., Kao, L.-J., Li, C., & Chu, C.-M. (2023). Integrating Structured and Unstructured EHR Data for Predicting Mortality by Machine Learning and Latent Dirichlet al.location Method. *International Journal of Environmental Research and Public Health*, 20(5), 4340. <https://doi.org/10.3390/ijerph20054340> Acesso em 16/01/2026.

CHO, K.-J., Kwon, O., Kwon, J., Lee, Y., Park, H., Jeon, K.-H., Kim, K.-H., Park, J., & Oh, B.-H. (2020). Detecting Patient Deterioration Using Artificial Intelligence in a Rapid Response System. *Critical Care Medicine*, 48(4), e285–e289. <https://doi.org/10.1097/CCM.0000000000004236> Acesso em 16/01/2026.

CHOI, A., CHOI, S. Y., CHUNG, K., CHUNG, H. S., SONG, T., CHOI, B., & KIM, J. H. (2023). Development of a machine learning- based clinical decision support system to predict clinical deterioration in patients visiting the emergency department. *Scientific Reports*, 13(1), 8561. <https://doi.org/10.1038/s41598-023-35617-3> Acesso em 16/01/2026.

CHOI, A., LEE, K., HYUN, H., KIM, K. J., AHN, B., LEE, K. H., HAHN, S., CHOI, S. Y., & KIM, J. H. (2024). A novel deep learning algorithm for real-time prediction of clinical deterioration in the emergency department for a multimodal clinical decision support system. *Scientific Reports*, 14(1), 30116. <https://doi.org/10.1038/s41598-024-80268-7> Acesso em 16/01/2026.

CONSIDINE, J., CASEY, P., OMONAIYE, O., VAN GULIK, N., ALLEN, J., & CURREY, J. (2024). Importance of specific vital signs in nurses' recognition and response to deteriorating patients: A scoping review. In *Journal of Clinical Nursing* (Vol. 33, Issue 7, pp. 2544–2561). John Wiley and Sons Inc. <https://doi.org/10.1111/jocn.17099> Acesso em 16/01/2026.

COUTINHO, J. G. F., ROCHA, R. G., ASSAD, L. G., ALMEIDA, L. F. DE, PAULA, V. G. DE, & LISBOA, C. DE D. (2025). Análise do uso de ferramentas de registro em prontuário eletrônico em Unidades de Terapia Intensiva. *Revista de Gestão e Secretariado*, 16(2), e4636. <https://doi.org/10.7769/gesec.v16i2.4636> Acesso em 16/01/2026.

- D'AGOSTINO SR., R. B., PENCINA, M. J., MASSARO, J. M., & COADY, S. (2013). Cardiovascular Disease Risk Assessment: Insights from Framingham. *Global Heart*, 8(1), 11. <https://doi.org/10.1016/j.gheart.2013.01.001> Acesso em 16/01/2026.
- DAS, P., AFROZ, F., RAHMAN, M. H., & SHABUZ, Z. R. (2025). Comparison of two measures of relative importance of predictors in logistic regression. *Discover Applied Sciences*, 7(4). <https://doi.org/10.1007/s42452-025-06818-4> Acesso em 16/01/2026.
- DATTA, S., MORASSI SASSO, A., KIWIT, N., BOSE, S., NADKARNI, G., MIOTTO, R., & BÖTTINGER, E. P. (2022). Predicting hypertension onset from longitudinal electronic health records with deep learning. *JAMIA Open*, 5(4). <https://doi.org/10.1093/jamiaopen/ooac097> Acesso em 16/01/2026.
- DEIMAZAR G, SHEIKHTAHERI A. Machine learning models to detect and predict patient safety events using electronic health records: A systematic review. *Int J Med Inform.* 2023 Dec;180:105246. <https://doi.org/10.1016/j.ijmedinf.2023.105246>. Epub 2023 Oct 9. PMID: 37837710. Acesso em 16/01/2026.
- DOSHI-VELEZ, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. <http://arxiv.org/abs/1702.08608> Acesso em 16/01/2026.
- Emergency Department at a Large Urban Hospital in the United States. *Journal of Emergency Medicine*, v. 58, n. 4, p. 570–580, 2020. Disponível em: < <https://doi.org/10.1016/j.jemermed.2019.11.020> >. Acesso em: 28/10/2025.
- ESCOBAR GJ, MYERS LC, LIU VX. Predictive Models for In-Hospital Deterioration in Ward Patients. *J Hosp Med.* 2021 Oct;16(10):640. doi: 10.12788/jhm.3675. <https://doi.org/10.12788/jhm.3675> Acesso em 16/01/2026.
- EVANGELISTA, J. R. G., SASSI, R. J., ROMERO, M., & NAPOLITANO, D. (2020). Systematic Literature Review to Investigate the Application of Open Source Intelligence (OSINT) with Artificial Intelligence. *Journal of Applied Security Research*, 16(3), 345–369. <https://doi.org/10.1080/19361610.2020.1761737> Acesso em 16/01/2026.
- FANELLI, S., PRATICI, L., SALVATORE, F. P., DONELLI, C. C., & ZANGRANDI, A. (2023). Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations. *Management Research Review*, 46(3), 369–389. <https://doi.org/10.1108/MRR-09-2021-0648> Acesso em 16/01/2026.
- FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH,P.; UTHURUSAMY, R. *Advances in KnowledgeDiscovery & Data Mining*. 1 ed. American Associationfor Artificial Intelligence, Menlo Park, Califórnia, 1996.611 folhas.
- FERNANDES PRABHU, D., GURUPUR, V., STONE, A., & TRADER, E. (2025). Integrating Artificial Intelligence, Electronic Health Records, and Wearables for Predictive, Patient-Centered Decision Support in Healthcare. *Healthcare*, 13(21), 2753. <https://doi.org/10.3390/healthcare13212753> Acesso em 16/01/2026.
- FRAILE NAVARRO, D., IJAZ, K., REZAZADEGAN, D., RAHIMI-ARDABILI, H., DRAS, M., COIERA, E., & BERKOVSKY, S. (2023). Clinical named entity recognition and relation extraction using natural language processing of medical free text: A systematic review. In *International Journal of Medical Informatics (Vol. 177)*. Elsevier Ireland Ltd. <https://doi.org/10.1016/j.ijmedinf.2023.105122> Acesso em 16/01/2026.
- FU LH, KNAPLUND C, CATO K, PEROTTE A, KANG MJ, DYKES PC, ALBERS D, COLLINS ROSSETTI S. Utilizing timestamps of longitudinal electronic health record data to classify clinical deterioration events. *J Am Med Inform Assoc.* 2021 Aug 13;28(9):1955-1963. <https://doi.org/10.1093/jamia/ocab111>. Acesso em 16/01/2026.
- N., Lim,. (2022). Classification of blood pressure levels based on photoplethysmogram and electrocardiogram signals with a concatenated convolutional neural network. *Diagnostics*, 12(11), 2886. <https://doi.org/10.3390/diagnostics12112886> Acesso em 16/01/2026.
- GIORGI, J. M., & BADER, G. D. (2018). Transfer learning for biomedical named entity recognition with neural networks. *Bioinformatics*, 34(23), 4087–4094. <https://doi.org/10.1093/bioinformatics/bty449> Acesso em 16/01/2026.
- GIORGI, J. M., & BADER, G. D. (2020). Towards reliable named entity recognition in the biomedical domain. *Bioinformatics*, 36(1), 280–286. <https://doi.org/10.1093/bioinformatics/btz504> Acesso em 16/01/2026.
- GONDIM, É. S., GOMES, E. B., MATOS, J. H. F. DE, PINTO, S. DE L., OLIVEIRA, C. J. DE, & ALENCAR, A. M. P. G. (2022). Tecnologias utilizadas pela enfermagem para predição de deterioração clínica em adultos hospitalizados: revisão de

escopo. *Revista Brasileira de Enfermagem*, 75(5). <https://doi.org/10.1590/0034-7167-2021-0570pt> Acesso em 16/01/2026.

GRAWE, M. F., & MOREIRA, V. P. (2023). Heterogeneous Ensemble Models for In-Hospital Mortality Prediction. *Anais Do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS 2023)*, 71–82. <https://doi.org/10.5753/sbcas.2023.229442> Acesso em 16/01/2026.

GROENHOF, T. K. J., et al (2020). Data mining information from electronic health records produced high yield and accuracy for current smoking status. *Journal of Clinical Epidemiology*, 118, 100–106. <https://doi.org/10.1016/j.jclinepi.2019.11.006> Acesso em 16/01/2026.

GUO J, DONG R, ZHANG R, YANG F, WANG Y, MIAO W. Interpretable machine learning model for predicting the prognosis of antibody positive autoimmune encephalitis patients. *J Affect Disord*. 2025 Jan 15;369:352-363. Epub 2024 Oct 5. DOI: [10.1016/j.jad.2024.10.010](https://doi.org/10.1016/j.jad.2024.10.010) Acesso em 16/01/2026.

GUPTA M, GALLAMOZA B, CUTRONA N, DHAKAL P, POULAIN R, BEHESHTI R. An Extensive Data Processing Pipeline for MIMIC-IV. *Proc Mach Learn Res*. 2022 Nov;193:311-325. <https://doi.org/10.48550/arXiv.2204.13841> Acesso em 16/01/2026.

HADDAWAY, N. R., PAGE, M. J., PRITCHARD, C. C., & MCGUINNESS, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, 18(2). <https://doi.org/10.1002/cl2.1230> Acesso em 16/01/2026.

HAMLIN, S. K., FONTENOT, N. M., HOOKER, S. J., & CHEN, H.-M. (2023). Systems-Based Physical Assessments: Earlier Detection of Clinical Deterioration and Reduced Mortality. *American Journal of Critical Care*, 32(5), 329–337. <https://doi.org/10.4037/ajcc2023113> Acesso em 16/01/2026.

HAQUE, S., EBERHART, Z., BANSAL, A., & MCMILLAN, C. (2022). Semantic Similarity Metrics for Evaluating Source Code Summarization. *IEEE International Conference on Program Comprehension*, 2022-March, 36–47. <https://doi.org/10.1145/nnnnnnn.nnnnnnn> Acesso em 16/01/2026.

HÄRKÄNEN, M., PAANANEN, J., MURRELLS, T., RAFFERTY, A. M., & FRANKLIN, B. D. (2019). Identifying risks areas related to medication administrations - text mining analysis using free-text descriptions of incident reports. *BMC Health Services Research*, 19(1), 791. <https://doi.org/10.1186/s12913-019-4597-9> Acesso em 16/01/2026

HARUTYUNYAN, H., KHACHATRIAN, H., KALE, D.C. et al. Multitask learning and benchmarking with clinical time series data. *Sci Data* 6, 96 (2019). <https://doi.org/10.1038/s41597-019-0103-9> Acesso em 16/01/2026.

HARUTYUNYAN, H., KHACHATRIAN, H., KALE, D. C., VER STEEG, G., & GALSTYAN, A. (2019). Multitask learning and benchmarking with clinical time series data. *Scientific Data*, 6, 96. <https://doi.org/10.1038/s41597-019-0103-9> Acesso em 16/01/2026.

HASHIR, M., & SAWHNEY, R. (2020). Towards unstructured mortality prediction with free-text clinical notes. *Journal of Biomedical Informatics*, 108. <https://doi.org/10.1016/j.jbi.2020.103489> Acesso em 16/01/2026.

HAYKIN, Simon. *Neural Networks and Learning Machines*. 3. ed. New York: Prentice Hall, 2009.

HE, K., MAO, R., LIN, Q., RUAN, Y., LAN, X., FENG, M., & CAMBRIA, E. (2023). A Survey of Large Language Models for Healthcare: from Data, Technology, and Applications to Accountability and Ethics. <https://doi.org/10.1016/j.inffus.2025.102963> Acesso em 16/01/2026.

HEAVNER, S. F., KUMAR, V. K., ANDERSON, W., AL-HAKIM, T., DASHER, P., ARMAIGNAC, D. L., CLERMONT, G., COBB, J. P., MANION, S., REMY, K. E., REUTER-RICE, K., & HAENDEL, M. (2024). Critical Data for Critical Care: A Primer on Leveraging Electronic Health Record Data for Research From Society of Critical Care Medicine’s Panel on Data Sharing and Harmonization. *Critical Care Explorations*, 6(11), e1179. <https://doi.org/10.1097/CCE.0000000000001179> Acesso em 16/01/2026.

HERRERO-ZAZO, M., FITZGERALD, T., TAYLOR, V., STREET, H., CHAUDHRY, A. N., BRADLEY, J. R., BIRNEY, E., & KEEVIL, V. L. (2023). Using machine learning to model older adult inpatient trajectories from electronic health records data. *IScience*, 26(1), 105876. <https://doi.org/10.1016/j.isci.2022.105876> Acesso em 16/01/2026.

HIMSS. (2021). *Preparing for Success in Healthcare Information and Management Systems*. HIMSS Publishing.

<https://doi.org/10.4324/9780367806163> Acesso em 16/01/2026.

HOLDSWORTH LM, KLING SMR, SMITH M, SAFAEINILI N, SHIEH L, VILENDRER S, GARVERT DW, WINGET M, ASCH SM, LI RC. Predicting and Responding to Clinical Deterioration in Hospitalized Patients by Using Artificial Intelligence: Protocol for a Mixed Methods, Stepped Wedge Study. *JMIR Res Protoc*. 2021 Jul 7;10(7):e27532. DOI: <https://doi.org/10.2196/27532> Acesso em 16/01/2026.

HOU, N., LI, M., HE, L., XIE, B., WANG, L., ZHANG, R., YU, Y., SUN, X., PAN, Z., & WANG, K. (2020). Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *Journal of Translational Medicine*, 18(1), 462. <https://doi.org/10.1186/s12967-020-02620-5> Acesso em 16/01/2026.

HU, X (2024) An algorithm strategy for precise patient monitoring in a connected healthcare enterprise. In *npj Digital Medicine* (Vol. 2, Issue 1). Nature Publishing Group. <https://doi.org/10.1038/s41746-019-0107-z> Acesso em 16/01/2026.

HU, Y.-H., HUNG, J.-H., HU, L.-Y., HUANG, S.-Y., & SHEN, C.-C. (2023). An analysis of Chinese nursing electronic medical records to predict violence in psychiatric inpatients using text mining and machine learning techniques. *PLOS ONE*, 18(6), e0286347. <https://doi.org/10.1371/journal.pone.0286347> Acesso em 16/01/2026.

HUANG, J., CAI, Y., WU, X., HUANG, X., LIU, J., & HU, D. (2024). Prediction of mortality events of patients with acute heart failure in intensive care unit based on deep neural network. *Computer Methods and Programs in Biomedicine*, 256, 108403. <https://doi.org/10.1016/j.cmpb.2024.108403> Acesso em 16/01/2026.

HUANG, M. S., LAI, P. T., LIN, P. Y., YOU, Y. T., TSAI, R. T. H., & HSU, W. L. (2020). Biomedical named entity recognition and linking datasets: Survey and our recent development. *Briefings in Bioinformatics*, 21(6), 2219–2238. <https://doi.org/10.1093/bib/bbaa054> Acesso em 16/01/2026.

J. KOBYLARZ RIBEIRO et al. (2020) "A Machine Learning Early Warning System: Multicenter Validation in Brazilian Hospitals," 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), Rochester, MN, USA, 2020, pp. 321-326, <https://doi.org/10.1109/CBMS49503.2020.00067> Acesso em 16/01/2026.

JALOLI, M., & CESCONE, M. (2023). Long-Term Prediction of Blood Glucose Levels in Type 1 Diabetes Using a CNN-LSTM-Based Deep Neural Network. *Journal of Diabetes Science and Technology*, 17(6), 1590–1601. <https://doi.org/10.1177/19322968221092785> Acesso em 16/01/2026.

JEONG, H. G., et al. (2021). Combined deep CNN–LSTM network-based multitasking learning architecture for noninvasive continuous blood pressure estimation using difference in ECG-PPG features. *Scientific Reports*, 11, Article 92997. <https://doi.org/10.1038/s41598-021-92997-0> Acesso em 16/01/2026.

JHF, M., & AMPG, A. (2022). Use of Technologies used by nursing to predict clinical deterioration in hospitalized adults: a scoping review Gondim ES, Gomes EB. In *Rev Bras Enferm* (Vol. 75, Issue 5).

JOHNSON AE, POLLARD TJ, SHEN L, LEHMAN LW, FENG M, GHASSEMI M, MOODY B, SZOLOVITS P, CELI LA, MARK RG. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016 May 24;3:160035. doi: 10.1038/sdata.2016.35. <https://doi.org/10.1038/sdata.2016.35> Acesso em 16/01/2026.

JOHNSON, A. E. W., BULGARELLI, L., SHEN, L., GAYLES, A., SHAMMOUT, A., HORNG, S., POLLARD, T. J., HAO, S., MOODY, B., GOW, B., LEHMAN, L. H., CELI, L. A., & MARK, R. G. (2023). MIMIC-IV, a freely accessible electronic health record dataset. *Scientific Data*, 10(1), 1. <https://doi.org/10.1038/s41597-022-01899-x> Acesso em 16/01/2026.

JOHNSON, A. E. W., POLLARD, T. J., SHEN, L., LEHMAN, L.-W. H., FENG, M., GHASSEMI, M., et al. (2016). MIMIC-III, a freely accessible critical care database. *Scientific Data*, 3, 160035. <https://doi.org/10.1038/sdata.2016.35>

JOHNSON, A., POLLARD, T., & MARK, R. (2016). Banco de Dados Clínicos MIMIC-III (versão 1.4). *PhysioNet*. <https://doi.org/10.13026/C2XW26> Acesso em 16/01/2026.

JONES, D., MITCHELL, I., HILLMAN, K., & STORY, D. (2013). Defining clinical deterioration. *Resuscitation*, 84(8), 1029–1034. <https://doi.org/10.1016/j.resuscitation.2013.01.013>

JURAEV, F., EL-SAPPAGH, S., ABDUKHAMIDOV, E., ALI, F., & ABUHMED, T. (2022). Multilayer dynamic ensemble model for intensive care unit mortality prediction of neonate patients. *Journal of Biomedical Informatics*, 135, 104216. <https://doi.org/10.1016/j.jbi.2022.104216> Acesso em 16/01/2026.

KAMDJE WABO, G., PRASSER, F., GIEREND, K., SIEGEL, F., & GANSLANDT, T. (2023). Data Quality– and Utility-Compliant Anonymization of Common Data Model–Harmonized Electronic Health Record Data: Protocol for a Scoping Review. *JMIR Research Protocols*, 12, e46471. <https://doi.org/10.2196/46471> Acesso em 16/01/2026.

KIA, A., TIMSINA, P., JOSHI, H. N., KLANG, E., GUPTA, R. R., FREEMAN, R. M., REICH, D. L., TOMLINSON, M. S., DUDLEY, J. T., KOHLI-SETH, R., MAZUMDAR, M., & LEVIN, M. A. (2020). MEWS++: Enhancing the Prediction of Clinical Deterioration in Admitted Patients through a Machine Learning Model. *Journal of Clinical Medicine*, 9(2), 343. <https://doi.org/10.3390/jcm9020343> Acesso em 16/01/2026.

KIPNIS P, TURK BJ, WULF DA, LAGUARDIA JC, LIU V, CHURPEK MM, ROMERO-BRUFU S, ESCOBAR GJ. Development and validation of an electronic medical record-based alert score for detection of inpatient deterioration outside the ICU. *J Biomed Inform*. 2016 Dec;64:10-19. Epub 2016 Sep 20. DOI: <https://doi.org/10.1016/j.jbi.2016.09.013> Acesso em 16/01/2026.

KOLASKI, K., LOGAN, L. R., & IOANNIDIS, J. P. A. (2023). Guidance to best tools and practices for systematic reviews. *Systematic Reviews*, 12(1). <https://doi.org/10.1186/s13643-023-02255-9> Acesso em 16/01/2026.

KOLECK, T. A., DREISBACH, C., BOURNE, P. E., & BAKKEN, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. *Journal of the American Medical Informatics Association*, 26(4), 364–379. <https://doi.org/10.1093/jamia/ocy173> Acesso em 16/01/2026.

KORACH, Z. T., YANG, J., ROSSETTI, S. C., CATO, K. D., KANG, M. J., KNAPLUND, C., SCHNOCK, K. O., GARCIA, J. P., JIA, H., SCHWARTZ, J. M., & ZHOU, L. (2020). Mining clinical phrases from nursing notes to discover risk factors of patient deterioration. *International Journal of Medical Informatics*, 135. <https://doi.org/10.1016/j.ijmedinf.2019.104053> Acesso em 16/01/2026.

KOTULA CA, MARTIN J, CAREY KA, EDELSON DP, DLIGACH D, MAYAMPURATH A, AFSHAR M, CHURPEK MM. Comparison of Multimodal Deep Learning Approaches for Predicting Clinical Deterioration in Ward Patients: Observational Cohort Study. *J Med Internet Res*. 2025 Jun 11;27:e75340. <https://doi.org/10.2196/75340> Acesso em 16/01/2026.

KOWSHER, M., TAHABILDER, A., ISLAM SANJID, M. Z., PROTTASHA, N. J., UDDIN, M. S., HOSSAIN, M. A., & KADER JILANI, M. A. (2021). LSTM-ANN & BiLSTM-ANN: Hybrid deep learning models for enhanced classification accuracy. *Procedia Computer Science*, 193, 131–140. <https://doi.org/10.1016/j.procs.2021.10.013> Acesso em 16/01/2026.

KRALJEVIC, Z., BEAN, D., SHEK, A., BENDAYAN, R., HEMINGWAY, H., YEUNG, J. A., DENG, A., BASTON, A., ROSS, J., IDOWU, E., TEO, J. T., & DOBSON, R. J. B. (2024). Foresight—a generative pretrained transformer for modelling of patient timelines using electronic health records: a retrospective modelling study. *The Lancet Digital Health*, 6(4), e281–e290. [https://doi.org/10.1016/S2589-7500\(24\)00025-6](https://doi.org/10.1016/S2589-7500(24)00025-6) Acesso em 16/01/2026.

LABKOFF, S., OLADIMEJI, B., KANNRY, J., SOLOMONIDES, A., LEFTWICH, R., KOSKI, E., JOSEPH, A. L., LOPEZ-GONZALEZ, M., FLEISHER, L. A., NOLEN, K., DUTTA, S., LEVY, D. R., PRICE, A., BARR, P. J., HRON, J. D., LIN, B., SRIVASTAVA, G., PASTOR, N., LUQUE, U. S., ... QUINTANA, Y. (2024). Toward a responsible future: recommendations for AI-enabled clinical decision support. *Journal of the American Medical Informatics Association*, 31(11), 2730–2739. <https://doi.org/10.1093/jamia/ocae209> Acesso em 16/01/2026.

LEE, J., YOON, W., KIM, S., KIM, D., KIM, S., SO, C. H., & KANG, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4), 1234–1240. <https://doi.org/10.1093/bioinformatics/btz682> Acesso em 16/01/2026.

LI, I., PAN, J., GOLDWASSER, J., VERMA, N., WONG, W. P., NUZUMLALI, M. Y., ROSAND, B., LI, Y., ZHANG, M., CHANG, D., TAYLOR, R. A., KRUMHOLZ, H. M., & RADEV, D. (2022). Neural Natural Language Processing for unstructured data in electronic health records: A review. In *Computer Science Review* (Vol. 46). Elsevier Ireland Ltd. <https://doi.org/10.1016/j.cosrev.2022.100511> Acesso em 16/01/2026.

LI, J., ZHOU, Y., JIANG, X., NATARAJAN, K., PAKHOMOV, S. V., LIU, H., & XU, H. (2021). Are synthetic clinical notes useful for real natural language processing tasks: A case study on clinical entity recognition. *Journal of the American Medical Informatics Association*, 28(10), 2193–2201. <https://doi.org/10.1093/jamia/ocab112> Acesso em 16/01/2026.

Lin, Y.-W., Zhou, Y., Faghri, F., Shaw, M. J., & Campbell, R. H. (2019). Analysis and prediction of unplanned intensive care unit readmission using recurrent neural networks with long short-term memory. *PLOS ONE*, 14(7), e0218942. <https://doi.org/10.1371/journal.pone.0218942> Acesso em 16/01/2026.

LIPTON, Z. C., Kale, D. C., Elkan, C., & Wetzell, R. (2015). Learning to Diagnose with LSTM Recurrent Neural Networks. <http://arxiv.org/abs/1511.03677> Acesso em 16/01/2026.

LOURENÇO, L. B. DE A., MESZAROS, M. DE J., SILVA, M. DE F. N., & SÃO-JOÃO, T. M. (2023). Nursing Training for Early Clinical Deterioration Risk Assessment: Protocol for an Implementation Study. *JMIR Research Protocols*, 12, e47293. <https://doi.org/10.2196/47293> Acesso em 16/01/2026.

LU, H., EHWERHEMUEPHA, L., & RAKOVSKI, C. (2022). A comparative study on deep learning models for text classification of unstructured medical notes with various levels of class imbalance. *BMC Medical Research Methodology*, 22(1), 181. <https://doi.org/10.1186/s12874-022-01665-y> Acesso em 16/01/2026.

LUO, Y., WANG, Z., & WANG, C. (2021). Improvement of APACHE II score system for disease severity based on XGBoost algorithm. *BMC Medical Informatics and Decision Making*, 21(1), 237. <https://doi.org/10.1186/s12911-021-01591-x> Acesso em 16/01/2026.

MAHBUB M, SRINIVASAN S, DANCIU I, PELUSO A, BEGOLI E, TAMANG S, PETERSON GD. Unstructured clinical notes within the 24 hours since admission predict short, mid & long-term mortality in adult ICU patients. *PLoS One*. 2022 Jan 6;17(1) DOI: [10.1371/journal.pone.0262182](https://doi.org/10.1371/journal.pone.0262182) Acesso em 16/01/2026.

MAHMOUD, A., & MOHAMMED, A. (2024). Leveraging Hybrid Deep Learning Models for Enhanced Multivariate Time Series Forecasting. *Neural Processing Letters*, 56(5). <https://doi.org/10.1007/s11063-024-11656-3>

MAKHNI, S., CERRATO, P., RICO, J., NIAZI, S., O'HORO, J., PETERS, S., SHAH, V., & HALAMKA, J. (2026). Meeting the challenges of electronic health record (EHR) optimization. In *npj Digital Medicine* (Vol. 9, Issue 1). Nature Research. <https://doi.org/10.1038/s41746-025-02178-w> Acesso em 16/01/2026.

MALM-NICOLAISEN, K., RUIZ, L. M., EVENSTAD, E. R., & PEDERSEN, R. (n.d.). Efforts on Using Standards for Defining the Structuring of Electronic Health Record Data: A Scoping Review.

MALYCHA, J, BACCHI, S, REDFERN, O. Artificial intelligence and clinical deterioration. *Curr Opin Crit Care*. 2022 Jun 1;28(3):315-321. <https://doi.org/10.1097/MCC.0000000000000945> Acesso em 16/01/2026.

MALYCHA, J., FCICM, M., MBBS, S. B., & REDFERN, O. (2022.). *Artificial Intelligence and Clinical Deterioration*.

MANN, K. D., GOOD, N. M., FATEHI, F., KHANNA, S., CAMPBELL, V., CONWAY, R., SULLIVAN, C., STAIB, A., JOYCE, C., & COOK, D. (2021). Predicting Patient Deterioration: A Review of Tools in the Digital Hospital Setting. *Journal of Medical Internet Research*, 23(9), e28209. <https://doi.org/10.2196/28209> Acesso em 16/01/2026.

MENG, C., Trinh, L., Xu, N., & Liu, Y. (2021). MIMIC-IF: Interpretability and Fairness Evaluation of Deep Learning Models on MIMIC-IV Dataset. <http://arxiv.org/abs/2102.06761> Acesso em 16/01/2026.

MENG, C., TRINH, L., XU, N., ENOUE, J., & LIU, Y. (2022). Interpretability and fairness evaluation of deep learning models on MIMIC-IV dataset. *Scientific Reports*, 12(1), 7166. <https://doi.org/10.1038/s41598-022-11012-2> Acesso em 16/01/2026.

MENG, C., TRINH, L., XU, N., & LIU, Y. (2021). MIMIC-IF: Interpretability and Fairness Evaluation of Deep Learning Models on MIMIC-IV Dataset. <http://arxiv.org/abs/2102.06761> Acesso em 16/01/2026.

MEYER A, ZVERINSKI D, PFAHRINGER B, KEMPFERT J, KUEHNE T, SÜNDERMANN SH, STAMM C, HOFMANN T, FALK V, EICKHOFF C. Machine learning for real-time prediction of complications in critical care: a retrospective study. *Lancet Respir Med*. 2018 Dec;6(12):905-914. Epub 2018 Sep 28. PMID: 30274956. [https://doi.org/10.1016/S2213-2600\(18\)30300-X](https://doi.org/10.1016/S2213-2600(18)30300-X) Acesso em 16/01/2026.

MIAO, F., WEN, B., HU, Z., FORTINO, G., WANG, X.-P., LIU, Z.-D., et al. (2020). Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques. *Artificial Intelligence in Medicine*, 108, 101919. <https://doi.org/10.1016/j.artmed.2020.101919> Acesso em 16/01/2026.

- MIOTTO, R., LI, L., KIDD, B. A., & DUDLEY, J. T. (2016). Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Scientific Reports*, 6. <https://doi.org/10.1038/srep26094> Acesso em 16/01/2026.
- MIOTTO, R., WANG, F., WANG, S., JIANG, X., & DUDLEY, J. T. (2018). Deep learning for healthcare: review, opportunities and challenges. *Briefings in Bioinformatics*, 19(6), 1236–1246. <https://doi.org/10.1093/bib/bbx044> Acesso em 16/01/2026.
- MISWAN, N. H., SULAIMAN, I. M., CHAN, C. S., & NG, C. G. (2021). Association rules mining for hospital readmission: A case study. *Mathematics*, 9(21). <https://doi.org/10.3390/math9212706> Acesso em 16/01/2026.
- MOHER, D., et al (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Espanola de Nutricion Humana y Dietetica*, 20(2), 148–160. <https://doi.org/10.1186/2046-4053-4-1> Acesso em 16/01/2026.
- MOORE, M., et al. (2021). Early prediction of sepsis in the ICU using machine learning (inclui CNN-LSTM como baseline em benchmarks recentes de UTI). *Frontiers in Medicine*, 8, 607952. <https://doi.org/10.3389/fmed.2021.607952> Acesso em 16/01/2026.
- MULLAINATHAN, S., & SPIESS, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106. <https://doi.org/10.1257/jep.31.2.87> Acesso em 16/01/2026.
- NA PATTALUNG T, INGVIYA T, CHAICHULEE S. Feature Explanations in Recurrent Neural Networks for Predicting Risk of Mortality in Intensive Care Patients. *J Pers Med*. 2021 Sep 19;11(9):934. <https://doi.org/10.3390/jpm11090934> Acesso em 16/01/2026
- NASARUDIN, N.A., AL JASMI, F., SINNOTT, R.O. et al. A review of deep learning models and online healthcare databases for electronic health records and their use for health prediction. *Artif Intell Rev* 57, 249 (2024). <https://doi.org/10.1007/s10462-024-10876-2> Acesso em 16/01/2026.
- NGUYEN, P. Y., et al (2022). Changing patterns in reporting and sharing of review data in systematic reviews with meta-analysis of the effects of interventions: Cross sectional meta-research study. *BMJ*. <https://doi.org/10.1136/bmj-2022-072428> Acesso em 16/01/2026.
- NING, Y., LI, S., NG, Y. Y., CHIA, M. Y. C., GAN, H. N., TIAH, L., MAO, D. R., NG, W. M., LEONG, B. S. H., DOCTOR, N., ONG, M. E. H., & LIU, N. (2024). Variable importance analysis with interpretable machine learning for fair risk prediction. *PLOS Digital Health*, 3(7). <https://doi.org/10.1371/journal.pdig.0000542> Acesso em 16/01/2026.
- OBERMEYER, Z., & EMANUEL, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, 375(13), 1216–1219. <https://doi.org/10.1056/nejmp1606181> Acesso em 16/01/2026.
- OECD/European Union/EC-JRC (2008), Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Publishing, Paris, <https://doi.org/10.1787/9789264043466-en>. Acesso em 16/01/2026.
- OLIVEIRA, R. R. DE, & DIAS, C. G. (2025). Uso de Aprendizado de Máquina em Registros Eletrônicos de Saúde para Prever Deteriorações Clínicas e Mortalidade em Pacientes Adultos Hospitalizados. *Revista de Gestão e Secretariado*, 16(4), e4882. <https://doi.org/10.7769/gesec.v16i4.4882> Acesso em 16/01/2026.
- PADILLA, R. M., & MAYO, A. M. (2018). Clinical deterioration: A concept analysis. In *Journal of Clinical Nursing* (Vol. 27, Issues 7–8, pp. 1360–1368). Blackwell Publishing Ltd. <https://doi.org/10.1111/jocn.14238> Acesso em 16/01/2026.
- PAGE, M. J., et al. (2021). PRISMA (2020) Updated guidance and exemplars for reporting systematic reviews. In *The BMJ* (Vol. 372). BMJ Publishing Group. <https://doi.org/10.1136/bmj.n160> Acesso em 16/01/2026.
- PEELEN RV, EDDAHCHOURI Y, KOENEMAN M, VAN DE BELT TH, VAN GOOR H, BREDIE SJ. Algorithms for Prediction of Clinical Deterioration on the General Wards: A Scoping Review. *J Hosp Med*. 2021 Oct;16(10):612-619. doi: 10.12788/jhm.3630. <https://doi.org/10.12788/jhm.3630> Acesso em 16/01/2026.
- PETERS, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis*, 18(10), 2119–2126. <https://doi.org/10.11124/JBIES-20-00167> Acesso em 16/01/2026.

PIRRACCHIO, R., COHEN, M. J., MALENICA, I., COHEN, J., CHAMBAZ, A., CANNESON, M., LEE, C., RESCHE-RIGON, M., & HUBBARD, A. (2019). Big data and targeted machine learning in action to assist medical decision in the ICU. *Anaesthesia Critical Care & Pain Medicine*, 38(4), 377–384. <https://doi.org/10.1016/j.accpm.2018.09.008> Acesso em 16/01/2026.

PURUSHOTHAM, S., MENG, C., CHE, Z., & LIU, Y. “Benchmark of Deep Learning Models on Large Healthcare MIMIC Datasets.” *Journal of Biomedical Informatics* Volume 83, July 2018, Pages 112-134 <https://doi.org/10.1016/j.jbi.2018.04.007> Acesso em 16/01/2026.

RAFIEI, M. H., & MELLO, M. M. (2021). Deep learning model for predicting sepsis using MIMIC-III. *Computers in Biology and Medicine*, 137, 104855. <https://doi.org/10.1016/j.compbimed.2021.104855> Acesso em 16/01/2026.

RAJKOMAR, A., et al (2018). Scalable and accurate deep learning with electronic health records. *Npj Digital Medicine*, 1(1), 18. <https://doi.org/10.1038/s41746-018-0029-1>

RIBEIRO SOARES, E., et al (2025). Aprendizagem de máquina na previsão de índices de mortalidade em Unidades de Terapia Intensiva: uma revisão integrativa da literatura. *Navus - Revista de Gestão e Tecnologia*, 16, 1–14. <https://doi.org/10.22279/navus.v16.2091> Acesso em 16/01/2026.

ROMERO-BRUFU S, WHITFORD D, JOHNSON MG, HICKMAN J, MORLAN BW, THERNEAU T, NAESSENS J, HUDDLESTON JM. Using machine learning to improve the accuracy of patient deterioration predictions: Mayo Clinic Early Warning Score (MC-EWS). *J Am Med Inform Assoc*. 2021 Jun 12;28(6):1207-1215. doi: 10.1093/jamia/ocaa347. <https://doi.org/10.1093/jamia/ocaa347> Acesso em 16/01/2026.

RUIZ, V. M., GOLDSMITH, M. P., SHI, L., SIMPAO, A. F., GÁLVEZ, J. A., NAIM, M. Y., NADKARNI, V., GAYNOR, J. W., & TSUI, F. (RICH). (2022). Early prediction of clinical deterioration using data-driven machine-learning modeling of electronic health records. *Journal of Thoracic and Cardiovascular Surgery*, 164(1), 211–222.e3. <https://doi.org/10.1016/j.jtcvs.2021.10.060> Acesso em 16/01/2026.

SANTOS, J. J. A. et al (2018). Adaptation and cross-cultural validation of the Brazilian version of the Warwick-Edinburgh mental well-being scale. *Revista Da Associação Médica Brasileira*, 61(3), 209–214. <https://doi.org/10.1590/1806-9282.61.03.209> Acesso em 16/01/2026.

SAVOVA, G. K., MASANZ, J. J., OGREN, P. V, ZHENG, J., SOHN, S., KIPPER-SCHULER, K. C., & CHUTE, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, 17(5), 507–513. <https://doi.org/10.1136/jamia.2009.001560> Acesso em 16/01/2026.

SHAH, P. K., GINESTRA, J. C., UNGAR, L. H., JUNKER, P., ROHRBACH, J. I., FISHMAN, N. O., & WEISSMAN, G. E. (2021). A Simulated Prospective Evaluation of a Deep Learning Model for Real-Time Prediction of Clinical Deterioration Among Ward Patients\*. *Critical Care Medicine*, 49(8), 1312–1321. <https://doi.org/10.1097/CCM.0000000000004966> Acesso em 16/01/2026.

SHAMOUT, F. E. (2023). Machine Learning for Decision Support Systems: Prediction of Clinical Deterioration (pp. 85–99). [https://doi.org/10.1007/978-3-031-17666-1\\_7](https://doi.org/10.1007/978-3-031-17666-1_7) Acesso em 16/01/2026.

SHAMOUT, F.E., ZHU T, SHARMA P, WATKINSON PJ, CLIFTON DA. Deep Interpretable Early Warning System for the Detection of Clinical Deterioration. *IEEE J Biomed Health Inform*. 2020 Feb;24(2):437-446. doi: 10.1109/JBHI.2019.2937803. Epub 2019 Sep 19. <https://doi.org/10.1109/JBHI.2019.2937803> Acesso em 16/01/2026.

SHEN, Y., YU, J., ZHOU, J., & HU, G. (2025). Twenty-Five Years of Evolution and Hurdles in Electronic Health Records and Interoperability in Medical Research: Comprehensive Review. *Journal of Medical Internet Research*, 27, e59024. <https://doi.org/10.2196/59024> Acesso em 16/01/2026.

SHICKEL, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589–1604. <https://doi.org/10.1109/JBHI.2017.2767063> Acesso em 16/01/2026.

SHRESTHA, A., ZIKOS, D., & FEGARAS, L. (2021). An annotated association mining approach for extracting and visualizing interesting clinical events. *International Journal of Medical Informatics*, 148, 104366. <https://doi.org/10.1016/j.ijmedinf.2020.104366> Acesso em 16/01/2026.

- SHU, X., & YE, Y. (2023). Knowledge Discovery: Methods from data mining and machine learning. *Social Science Research*, 110. <https://doi.org/10.1016/j.ssresearch.2022.102817> Acesso em 16/01/2026.
- SILVA, G. F. DOS S., WICHMANN, R. M., DA SILVA JUNIOR, F. C., & CHIAVEGATTO FILHO, A. D. P. (2025). Development and evaluation of machine learning training strategies for neonatal mortality prediction using multicountry data. *Scientific Reports*, 15(1), 24278. <https://doi.org/10.1038/s41598-025-04066-5> Acesso em 16/01/2026.
- SINGH, D., & SINGH, B. (2020). Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, 97, 105524. <https://doi.org/10.1016/j.asoc.2019.105524> Acesso em 16/01/2026.
- SLATER, L. T., BRADLOW, W., BALL, S., HOEHNDORF, R., & GKOUTOS, G. V. (2021). Improved characterisation of clinical text through ontology-based vocabulary expansion. *Journal of Biomedical Semantics*, 12(1). <https://doi.org/10.1186/s13326-021-00241-5> Acesso em 16/01/2026.
- STEINBERG, E., JUNG, K., FRIES, J. A., CORBIN, C. K., PFOHL, S. R., & SHAH, N. H. (2021). Language models are an effective representation learning technique for electronic health record data. *Journal of Biomedical Informatics*, 113, 103637. <https://doi.org/10.1016/j.jbi.2020.103637> Acesso em 16/01/2026.
- STEYERBERG, E. W. "FRANK E. HARRELL, (2018) (Regression Modeling Strategies: With Applications, to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Heidelberg: Springer." (2018): 1006-1007.
- SULLIVAN, L. M., MASSARO, J. M., & D'AGOSTINO, R. B. (2024). Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in Medicine*, 23(10), 1631–1660. <https://doi.org/10.1002/sim.1742> Acesso em 16/01/2026.
- SUN, W., CAI, Z., LI, Y., LIU, F., FANG, S., & WANG, G. (2018). Data processing and text mining technologies on electronic medical records: A review. In *Journal of Healthcare Engineering* (Vol. 2018). Hindawi Limited. <https://doi.org/10.1155/2018/4302425> Acesso em 16/01/2026.
- SUNDERMANN, A. J., MILLER, J. K., MARSH, J. W., SAUL, M. I., SHUTT, K. A., PACEY, M., MUSTAPHA, M. M., AYRES, A., PASCULLE, A. W., CHEN, J., SNYDER, G. M., DUBRAWSKI, A. W., & HARRISON, L. H. (2019). Automated data mining of the electronic health record for investigation of healthcare-associated outbreaks. *Infection Control & Hospital Epidemiology*, 40(3), 314–319. <https://doi.org/10.1017/ice.2018.343> Acesso em 16/01/2026.
- TAENZER, A.H., SPENCE, B.C. (2018) Continuous Monitoring on General Care Units. *Apr*;34(2):189-198. doi: 10.1016/j.ccc.2017.12.001. Acesso em 26/01/2026.
- TAMMINEN J, KALLONEN A, HOPPU S, KALLIOMÄKI J. Machine learning model predicts short-term mortality among prehospital patients: A prospective development study from Finland. *Resusc Plus*. 2021 Feb 5;5:100089. DOI: [10.1016/j.resplu.2021.100089](https://doi.org/10.1016/j.resplu.2021.100089) Acesso em 16/01/2026.
- TANG, S., et al (2020). Democratizing EHR analyses with FIDDLE: A flexible data-driven preprocessing pipeline for structured clinical data. *Journal of the American Medical Informatics Association*, 27(12), 1921–1934. <https://doi.org/10.1093/jamia/ocaa139> Acesso em 16/01/2026.
- TIAN, J., XIANG, A., FENG, Y., YANG, Q., & LIU, H. (2024). Enhancing Disease Prediction with a Hybrid CNN-LSTM Framework in EHRs. *Journal of Theory and Practice of Engineering Science*, 4(02), 8–14. [https://doi.org/10.53469/jtpes.2024.04\(02\).02](https://doi.org/10.53469/jtpes.2024.04(02).02) Acesso em 16/01/2026.
- TIBSHIRANI, R. (1996). Regression Shrinkage and Selection via the Lasso. In *Source: Journal of the Royal Statistical Society. Series B (Methodological)* (Vol. 58, Issue 1).
- VAN DIJK, W. B., FIOLET, A. T. L., SCHUIT, E., SAMMANI, A., GROENHOF, T. K. J., VAN DER GRAAF, R., DE VRIES, M. C., ALINGS, M., SCHAAP, J., ASSELBERGS, F. W., GROBBEE, D. E., GROENWOLD, R. H. H., & MOSTERD, A. (2021). Text-mining in electronic healthcare records can be used as efficient tool for screening and data collection in cardiovascular trials: a multicenter validation study. *Journal of Clinical Epidemiology*, 132, 97–105. <https://doi.org/10.1016/j.jclinepi.2020.11.014> Acesso em 16/01/2026.
- VEGT, A. H. V. D., et al (2024). Systematic review and longitudinal analysis of implementing Artificial Intelligence to predict clinical deterioration in adult hospitals: what is known and what remains uncertain, *Journal of the American Medical Informatics Association*, Volume 31, Issue 2, February 2024, Pages 509–524, <https://doi.org/10.1093/jamia/ocad220> Acesso em 16/01/2026.

VELDHUIS, L. I., WOITTEZ, N. J. C., NANAYAKKARA, P. W. B., & LUDIKHUIZE, J. (2022). Artificial Intelligence for the Prediction of In-Hospital Clinical Deterioration: A Systematic Review. *Critical Care Explorations*, 4(9), e0744. <https://doi.org/10.1097/CCE.0000000000000744> Acesso em 16/01/2026.

VICENT, J.L., MORENO, R., TAKALA, J. et al (1996). The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* 22, 707–710. <https://doi.org/10.1007/BF01709751>. Acesso em 26/01/2026.

VUOKKO, R., VAKKURI, A., & PALOJOKI, S. (2023). Systematized Nomenclature of Medicine–Clinical Terminology (SNOMED CT) Clinical Use Cases in the Context of Electronic Health Record Systems: Systematic Literature Review. *JMIR Medical Informatics*, 11, e43750. <https://doi.org/10.2196/43750> Acesso em 16/01/2026.

Wang, W., Mohseni, P., Kilgore, K. L., & Najafizadeh, L. (2023). PulseDB: A large, cleaned dataset based on MIMIC-III and VitalDB for benchmarking cuff-less blood pressure estimation methods. *Frontiers in Digital Health*, 4, 1090854. <https://doi.org/10.3389/fgth.2022.1090854> Acesso em 16/01/2026.

WANG, X., ZHANG, Y., REN, X., ZHANG, Y., ZITNIK, M., SHANG, J., LANGLOTZ, C., & HAN, J. (2019). Cross-type biomedical named entity recognition with deep multi-task learning. *Bioinformatics*, 35(10), 1745–1752. <https://doi.org/10.1093/bioinformatics/bty869> Acesso em 16/01/2026.

YANG, F., WANG, X., MA, H., & LI, J. (2021). Transformers-sklearn: a toolkit for medical language understanding with transformer-based models. *BMC Medical Informatics and Decision Making*, 21. <https://doi.org/10.1186/s12911-021-01459-0> Acesso em 16/01/2026.

YANG, H., KUANG, L., & XIA, F. (2021). Multimodal temporal-clinical note network for mortality prediction. *Journal of Biomedical Semantics*, 12, 3. <https://doi.org/10.1186/s13326-021-00235-3> Acesso em 16/01/2026.

YAZDANPANA, H., SILVA, A. C. M., GUEDES, M., MORALES, H. M. P., DOS, L., COELHO, S., & MORO, F. G. (2024). clinical deterioration prediction in brazilian hospitals based on artificial neural networks and tree decision models. <https://doi.org/10.48550/arXiv.2212.08975> Acesso em 16/01/2026.

ZAMMEL, Z., KHABOU, N., SOUFI, L., & BOUASSIDA RODRIGUEZ, I. (2024). *Time Series Prediction Models in Healthcare: Systematic Literature Review*. 1286–1293. <https://doi.org/10.5220/0012465000003636> Acesso em 02/02/2026

ZHANG, D., YIN, C., ZENG, J., YUAN, X., & ZHANG, P. (2020). Combining structured and unstructured data for predictive models: a deep learning approach. *BMC Medical Informatics and Decision Making*, 20(1), 280. <https://doi.org/10.1186/s12911-020-01297-6> Acesso em 16/01/2026.

ZHANG, Y., ZHANG, Y., QI, P., MANNING, C. D., & LANGLOTZ, C. P. (2021). Biomedical and clinical English model packages for the Stanza Python PNL library. *Journal of the American Medical Informatics Association*, 28(9), 1892–1899. <https://doi.org/10.1093/jamia/ocab090> Acesso em 16/01/2026.

ZHUANG, Y., ZHANG, J., LI, X., LIU, C., YU, Y., DONG, W., & HE, K. (2025). Autonomous International Classification of Diseases Coding Using Pretrained Language Models and Advanced Prompt Learning Techniques: Evaluation of an Automated Analysis System Using Medical Text. *JMIR Medical Informatics*, 13, e63020. <https://doi.org/10.2196/63020> Acesso em 16/01/2026.