



Programa de Pós-Graduação em
Computação Aplicada

Doutorado Acadêmico

Luis Claudio Gubert

AIDA: Uma arquitetura inteligente para gerenciamento de diabetes

São Leopoldo, 2024

Luis Claudio Gubert

**AIDA: UMA ARQUITETURA INTELIGENTE PARA GERENCIAMENTO DE
DIABETES**

Tese apresentada como requisito parcial para a
obtenção do título de Doutor pelo Programa de
Pós-Graduação em Computação Aplicada da
Universidade do Vale do Rio dos Sinos —
UNISINOS

Orientador:
Prof. Dr. Cristiano André da Costa

Coorientador:
Prof. Dr. Rafael Kunst

São Leopoldo
2024

G921a Gubert, Luis Claudio
AIDA : uma arquitetura inteligente para gerenciamento
de diabetes / por Luis Claudio Gubert. – 2024.
168 f. : il.; 30 cm.

Tese (doutorado) — Universidade do Vale do Rio dos
Sinos, Programa de Pós-Graduação em Computação
Aplicada, 2024.

Orientador: Prof. Dr. Cristiano André da Costa ;
Coorientador: Prof. Dr. Rafael Kunst.

1. Ciência de contexto. 2. IoT. 3. Aprendizado de
máquina. 4. Diabetes mellitus. 5. Cuidados com saúde.
I. Título.

CDU 004.738.5

Catálogo na Fonte:
Bibliotecária Vanessa Borges Nunes - CRB 10/1556

ATA DE BANCA EXAMINADORA DE TESE DE DOUTORADO Nº 01/2024

Aluno: Luis Claudio Gubert

Título da Tese: "AIDA: UMA ARQUITETURA INTELIGENTE PARA GERENCIAMENTO DE DIABETES"

Banca: Prof. Dr. Cristiano André da Costa (Orientador) - Unisinos
Prof. Dr. Rafael Kunst (Coorientador) - Unisinos
Prof. Dr. José Palazzo Moreira de Oliveira (Avaliador) - UFRGS
Profa. Dra. Marta Rosecler Bez (Avaliadora) - FEEVALE
Profa. Dr. Sandro Jose Rigo (Avaliador) – Unisinos

Ao primeiro dia do mês de abril do ano de 2024, às 14h reuniu-se a Comissão Examinadora de Defesa de Tese composta pelos professores: Prof. Dr. Cristiano André da Costa (Orientador) – Unisinos (por webconferência); Prof. Dr. Rafael Kunst (Coorientador) – Unisinos (por webconferência); Prof. Dr. José Palazzo Moreira de Oliveira (Avaliador) – UFRGS (por webconferência); Profa. Dra. Marta Rosecler Bez (Avaliadora) – FEEVALE e Profa. Dr. Sandro Jose Rigo (Avaliador) – Unisinos (por webconferência) para analisar e avaliar a Tese apresentada pelo(a) aluno(a) **Luis Claudio Gubert** (por webconferência).

Considerações da Banca:

Após deliberação, os membros da banca avaliadora consideraram que o trabalho atende aos critérios esperados para uma tese de doutorado, além de apresentar originalidade científica na área de pesquisa. Mereceu destaque pela banca o trabalho desde o aporte teórico até os experimentos realizados. A banca sugeriu algumas melhorias para o trabalho que devem ser incorporadas na versão final do documento.

Ocorreu alteração do título? (X) Não () Sim
Indicar o novo título:

A Banca Examinadora, em cumprimento ao requisito exigido para a obtenção do Título de Doutor em Computação Aplicada, julga esta tese:

(X) APROVADA () REPROVADA

Conforme Artigo 75 do Regimento do Programa o texto definitivo, com aprovação do Orientador, deverá ser entregue no prazo máximo de sessenta (60) dias após a defesa. O resultado da banca é de consenso entre os avaliadores. A emissão do Diploma está condicionada a entrega da versão final da Tese.

São Leopoldo, 01 de abril de 2024.



Orientador - Prof. Dr. Cristiano André da Costa

*"Dedico este trabalho à minha esposa, Lisiane e a minha filha Maria Luísa, pela compreensão e paciência durante todo o período.
Aos meus pais José e Nadir, pelo incentivo em forma de exemplo, que me apresentam desde a mais tenra idade.
Sem o apoio de vocês eu não teria conseguido."*

AGRADECIMENTOS

À minha família, em especial a minha esposa, pelas contribuições e conversas importantes sobre a estrutura do trabalho e dos artigos.

Ao Dr. Cristiano André da Costa meu orientador, pela orientação, motivação e incentivo durante o percurso da Tese.

Ao Dr. Rafael Kunst, meu coorientador, pelas dicas e contribuições assertivas junto a temática desenvolvida neste trabalho.

Aos colegas *consagrados* do *SoftwareLab*, Felipe André Zeiser, Guilherme Goldschmidt e João Montenegro, que se tornaram amigos pessoais.

Aos professores e funcionários do PPGCA que sempre foram solícitos as questões levantadas e deram sua sábia e valorosa contribuição.

Ao Instituto Federal de Educação, Ciência e Tecnologia do Estado do Rio Grande do Sul - IFRS, pelo apoio financeiro junto a realização deste doutorado.

À Dra. Ângela Paveglio Teixeira Farias, médica endocrinologista e professora da Univates, pelas dicas e pelo conhecimento repassado acerca de diabetes e doenças crônicas, contribuindo significativamente no planejamento do escopo deste trabalho.

Ao Centro Clínico Univates por abrir as portas para este trabalho, contribuindo significativamente nos resultados deste trabalho.

A todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

RESUMO

A incidência do diabetes mellitus tipo 2 tem aumentado significativamente nos últimos anos, se destacando como uma causa significativa de morbidade e mortalidade. Atualmente o sistema de gerenciamento de doenças se concentra na identificação de uma doença e seus métodos curáveis. Essa abordagem não é adequada para pacientes com DM2, pois a doença é crônica e a abordagem deve se preocupar com o longo curso da doença e com todos os envolvidos no seu acompanhamento (o paciente, o serviço de atenção primária, médicos e hospitais). A evolução dos dispositivos portáteis e sensores, e o aumento do uso do registro eletrônico de saúde, contribuem para o aumento do volume de dados que podem ser utilizados para acompanhamento, melhoria e individualização do tratamento. A grande disponibilidade de dados para cada paciente torna desafiador o processo de análise pelos serviços de saúde. Tal fato, impulsiona a utilização de técnicas de inteligência artificial (IA) para extrair conhecimento que possa ser utilizado como suporte às decisões dos profissionais de saúde. Tendo em vista que cuidados de doenças crônicas são de longo curso e exigem um acompanhamento contínuo, propomos como contribuição deste trabalho o desenvolvimento de uma arquitetura computacional que, baseada nos dados coletados por sensores e dados do registro eletrônico de saúde do paciente, utilize o aprendizado de máquina para encontrar padrões individualizados do curso da doença. Pretende-se com isso, detectar precocemente o surgimento de comorbidades e o conseqüente declínio da saúde do paciente. Para esta tese propomos o desenvolvimento, implementação e avaliação de uma arquitetura inteligente para acompanhamento de pacientes com diabetes mellitus tipo 2, denominada AIDA. Avaliamos o modelo por meio de estudos de avaliação de aceitação da tecnologia baseados no TAM (technology acceptance model) e da aplicação do SUS (System Usability scale) em entrevistas com especialistas e pelo uso da tecnologia por parte de pacientes. Com relação às avaliações de classificação e predição, construímos um dataset com dados de um centro clínico nacional. Para avaliação dos resultados, foram comparados vários modelos de aprendizado de máquina para classificação, como *Random forest*, *Decision Tree*, *Logistic regression*, *Gradient boosting*, *XGBoost* e *LightGBM* com as configurações de hiperparâmetros alteradas para obter os melhores resultados. Para o modelo de predição foi utilizada redes LSTM (*Long short-term memory*) com dados coletados no contexto do paciente. Os resultados foram avaliados de forma visual e por métricas como a acurácia, precisão, sensibilidade, especificidade, F1-score e AUROC para a classificação e *Root mean square error* (RMSE), *Mean absolute error* (MAE) e *Mean Poisson Deviance* (MPD) para predição. Os resultados de aceitação e usabilidade apontaram que ambos os grupos, profissionais e pacientes, tem percepções positivas sobre a aplicação de tecnologia nos cuidados de doenças crônicas. Por sua vez, os melhores resultados para classificação apresentaram AUC de 0.85, especificidade de 0.92, sensibilidade de 0.52, precisão de 0.46, acurácia de 0.87 e *F1-score* de 0.49. Para predição, os melhores resultados apresentaram RMSE de 38.74, MAE de 31.41 e MPD de 8.82. Os resultados obtidos reforçam a hipótese de que é possível definir um modelo computacional para suporte e acompanhamento de pacientes com diabetes mellitus tipo 2.

Palavras-chave: Ciência de contexto. IoT. Aprendizado de máquina. Diabetes mellitus. Cuidados com saúde.

ABSTRACT

The incidence of type 2 diabetes mellitus has increased significantly in recent years, emerging as a significant cause of morbidity and mortality. The disease management system currently focuses on identifying a disease and its curable methods. This approach is not suitable for patients with DM2, as the disease is chronic, and the approach must be concerned with the long course of the disease and with everyone involved in its follow-up (the patient, the primary care service, doctors, and hospitals). The evolution of portable devices and sensors and the increased use of electronic health records contribute to the increase in the volume of data that can be used to monitor, improve, and individualize treatment. The extensive data availability for each patient makes the analysis process by health services challenging. This fact encourages the use of artificial intelligence (AI) techniques to extract knowledge that can be used to support healthcare professionals' decisions. Considering that care for chronic diseases is long-term. It requires continuous monitoring, we propose as a contribution to this work the development of a computational architecture that, based on data collected by sensors and data from the patient's electronic health record, uses machine learning to find individualized patterns of the course of the disease. The aim is to detect the emergence of comorbidities early and the consequent decline in the patient's health. For this thesis, we propose developing, implementing, and evaluating an intelligent architecture called AIDA for monitoring patients with type 2 diabetes mellitus. We evaluated the model through technology acceptance assessment studies based on the TAM (technology acceptance model) and the application of the SUS (System Usability scale) in interviews with experts and through the use of technology by patients. Regarding classification and prediction evaluations, we built a dataset with data from a national clinical center. To evaluate the results, several machine learning models for classification were compared, such as *Random forest*, *Decision Tree*, *Logistic regression*, *Gradient boosting*, *XGBoost* and *LightGBM* with hyperparameter settings changed to get the best results. LSTM networks (*Long short-term memory*) were used for the prediction model with data collected in the patient's context. The results were evaluated visually and by metrics such as accuracy, precision, sensitivity, specificity, F1-score, and AUROC for classification and *Root mean square error* (RMSE), *Mean absolute error* (MAE) and *Mean Poisson Deviance* (MPD) for prediction. The acceptance and usability results showed that both groups, professionals and patients, have positive perceptions about applying technology in the care of chronic diseases. In turn, the best results for classification presented AUC of 0.85, specificity of 0.92, sensitivity of 0.52, precision of 0.46, accuracy of 0.87 and *F1-score* of 0.49. For prediction, the best results presented RMSE of 38.74, MAE of 31.41, and MPD of 8.82. The results reinforce the hypothesis that it is possible to define a computational model to support and monitor patients with type 2 diabetes mellitus.

Keywords: Context-aware. IoT. Machine learning. Diabetes mellitus. Healthcare.

LISTA DE FIGURAS

Figura 1 – Sistema de Acompanhamento Inteligente	24
Figura 2 – Proporções de números de óbitos de acordo com as causas.	30
Figura 3 – Proporção de mortes causadas pela DM2 no Brasil	31
Figura 4 – Modelo para melhoria do tratamento de doenças crônicas	33
Figura 5 – Fluxograma de tratamento de pacientes com DM2	34
Figura 6 – Taxonomia de técnicas de aprendizagem supervisionada	39
Figura 7 – Taxonomia de técnicas de aprendizagem não supervisionada	40
Figura 8 – Taxonomia das técnicas de aprendizagem por reforço	41
Figura 9 – Classificação de pacientes com alto e baixo risco de gripe	43
Figura 10 – LSTM tradicional	45
Figura 11 – Aprendizado de máquina online versus aprendizado em lote. (a) fluxo de trabalho de aprendizado de máquina em lote; (b) Fluxo de trabalho de aprendizado de máquina online.	46
Figura 12 – Artigos selecionados por banco de dados	55
Figura 13 – Número de artigos por ano de publicação	58
Figura 14 – Número de citações por artigo	59
Figura 15 – Taxonomia do uso de ciência de contexto em saúde	60
Figura 16 – Resumo dos desafios	69
Figura 17 – Tecnologias AIDA	76
Figura 18 – Visão geral do modelo AIDA	77
Figura 19 – Arquitetura AIDA com seus módulos e submódulos	78
Figura 20 – Algoritmo para filtrar valores não viáveis	80
Figura 21 – Algoritmo para filtrar valores faltantes e <i>outliers</i>	81
Figura 22 – Submódulo de classificação	83
Figura 23 – Submódulo de regressão	83
Figura 24 – Fluxo dos Experimentos realizados	89
Figura 25 – Visão geral dos dados utilizados	91
Figura 26 – Aplicativo AIDA	95
Figura 27 – Gráfico com medições de glicose capilar e linha de tendência	99
Figura 28 – Gráficos das técnicas utilizadas para imputação de dados.	101
Figura 29 – Técnicas de Balanceamento de dados utilizadas	102
Figura 30 – Modelo de Aceitação de Tecnologia	103
Figura 31 – Divisão do conjunto de dados	107
Figura 32 – Seleção do Modelo	108
Figura 33 – Grid de Erro de Clarke	113
Figura 34 – Análise dos coeficientes de correlação de Pearson entre os itens dos construtos	117
Figura 35 – Mapa de calor de 10 itens avaliando a aceitação de tecnologia de 17 usuários usando 3 construtos do Modelo de Aceitação de Tecnologia	119
Figura 36 – Fluxo de Experimentos de Classificação Realizados	120
Figura 37 – Melhores AUROC dos Modelos	126
Figura 38 – Importância das Variáveis Preditoras	127
Figura 39 – Gráficos de Predição com as técnicas utilizadas para imputação de dados.	129
Figura 40 – Erro de Grid de Clarke para cada Tipo de Imputação de Dados.	131

LISTA DE TABELAS

Tabela 1 –	Intervalos de referência para sinais vitais e laboratoriais em adultos diabéticos	36
Tabela 2 –	Questões de pesquisa	52
Tabela 3 –	Critérios de avaliação da qualidade	54
Tabela 4 –	Revisão dos artigos relacionados às perguntas da pesquisa	54
Tabela 5 –	Corpus final de artigos publicados	56
Tabela 6 –	Desafios	62
Tabela 7 –	Cenários	64
Tabela 8 –	Métodos de fusão e raciocínio	68
Tabela 9 –	Corpus de artigos Diabetes Mellitus tipo 2, machine learning e Predição . .	72
Tabela 10 –	Tabela de possíveis decisões do submódulo Ciência de contexto, baseada nos resultados do módulo Aprendizado de máquina	85
Tabela 11 –	Fotografia temporal de dados coletados no sistema do centro clínico. . . .	92
Tabela 12 –	Exemplo de dados no campo evolução, coletados no sistema do centro clínico.	93
Tabela 13 –	Detalhamento das variáveis utilizadas para classificação	94
Tabela 14 –	Dados de glicose capilar coletados via AIDA app	96
Tabela 15 –	Dados de peso coletados via AIDA app	96
Tabela 16 –	Dados de temperatura coletados via AIDA app	97
Tabela 17 –	Leituras de pressão arterial e batimentos cardíacos	97
Tabela 18 –	Tabela com quantidade e porcentagem de dados faltantes	98
Tabela 19 –	Itens dos Construtos	105
Tabela 20 –	Questionário baseado nas afirmações do SUS	106
Tabela 21 –	Índice de validade de conteúdo por item, por escala e alfa de cronbach dos construtos	116
Tabela 22 –	Informação Demográfica	118
Tabela 23 –	Características demográficas e clínicas dos pacientes	122
Tabela 24 –	Características clínicas dos pacientes - Adicção	123
Tabela 25 –	Características clínicas dos pacientes - Medicções	124
Tabela 26 –	Performance dos Modelos de Aprendizado de Máquina (melhores AUROC)	125
Tabela 27 –	Matriz de Confusão dos algoritmos (melhores AUROC)	125
Tabela 28 –	Performance dos Modelos de Aprendizado de Máquina (melhores <i>F1-score</i>)	126
Tabela 29 –	Matriz de Confusão dos melhores algoritmos (melhores <i>F1-score</i>)	127
Tabela 30 –	Resultados dos Critérios de Avaliação por Tipo de Imputação de Dados . .	130
Tabela 31 –	Porcentagem de Pontos de Predição nas Zonas do Grid de Erro de Clarke. .	130
Tabela 32 –	Comparação do modelo de classificação AIDA com trabalhos relacionados.	137

LISTA DE SIGLAS

AUC	Area Under the Curve
ASP	Programação de Conjunto de Resposta
BD	Banco de Dados
DCNT	Doença Crônica não Transmissível
DM	Diabetes mellitus
DM2	Diabetes mellitus tipo 2
ES	Sistemas especialistas
ANN	Artificial Neural Network
AUC	Area Under the ROC curve
CNN	Convolutional Neural Networks
DL	Deep Learning
FL	Lógica Fuzzy
FN	Falso Negativo
FP	Falso Positivo
IA	Inteligencia Artificial
IQR	Intervalo Inter-Quartil
IMC	Índice de Massa Corporal
IoT	Internet das Coisas
KNN	k-Nearest Neighbors
LSTM	Long Short-Term Memory
OMS	Organização Mundial da Saúde
OWL	Ontology Web Language
PA	Pressão Arterial
PNL	Processamento de Linguagem Natural
RDF	Resource Description Framework
RES	Registro Eletrônico de Saúde
RMSE	Root Mean Squared Error
RNN	Redes Neurais Recorrentes
ROC	Receive Operating characteristic
RSL	Revisão Sistemática de Literatura
SUS	Sistema Único de Saúde
SVM	Support Vector Machines
TPF	Taxa Positiva Falsa

TPV	Taxa Positiva Verdadeira
UML	Linguagem de Modelagem Unificada
XML	Extensible Markup Language
ROC	Receiver Operating Characteristic
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Motivação	23
1.2	Problema de pesquisa	24
1.3	Organização do documento	27
2	FUNDAMENTAÇÃO TEÓRICA	29
2.1	Diabetes melittus tipo 2	29
2.1.1	Tratamento da diabetes mellitus tipo 2	31
2.1.2	Sinais vitais e laboratoriais para monitoramento de DM2	35
2.2	Inteligência artificial	37
2.2.1	Tipos de aprendizado	38
2.2.2	Objetivos na aprendizagem de máquina	41
2.2.3	LSTM	44
2.2.4	Aprendizagem de máquina online	45
2.3	Considerações parciais	47
3	TRABALHOS RELACIONADOS	49
3.1	Método	51
3.1.1	Projeto de estudo	51
3.1.2	Questões de pesquisa	51
3.1.3	Estratégia de pesquisa	52
3.1.4	Seleção de Artigos	53
3.1.5	Avaliação da qualidade	53
3.1.6	Extração de dados	54
3.2	Resultados	54
3.2.1	Estratégia para conduzir a pesquisa	55
3.2.2	Prosseguindo com a seleção de artigos	55
3.2.3	Avaliando a qualidade	57
3.3	Respostas às perguntas e discussão	57
3.3.1	Como o uso da ciência de contexto no campo da saúde é definido e classificado nas pesquisas existentes?	58
3.3.2	Quais são os desafios relacionados ao uso de informações de ciência de contexto na área da saúde?	61
3.3.3	Quais são os cenários relacionados à ciência de contexto na área da saúde?	64
3.3.4	Quais são as técnicas de modelagem de contexto empregadas na área da saúde?	66
3.3.5	Quais são os métodos de raciocínio e fusão de dados usados em informações de ciência de contexto em saúde?	67
3.4	Desafios e direções futuras	68
3.5	Ciência de contexto e <i>diabetes mellitus</i>	70
3.6	Considerações Parciais	72
4	MODELO AIDA	75
4.1	Decisões de projeto	75
4.2	Visão geral	77
4.3	Detalhamento do modelo	77

4.3.1	Módulo Tecnologias de Suporte	79
4.3.2	Módulo Coletor de dados	79
4.3.3	Módulo Pré-processamento	80
4.3.4	Módulo Aprendizado de máquina	82
4.3.5	Módulo Decisão	84
4.4	Considerações parciais	87
5	MATERIAIS E MÉTODOS	89
5.1	Declaração de Ética	90
5.2	Materiais	90
5.2.1	Caracterização Centro Clínico	90
5.2.2	Dados empregados	91
5.2.3	Pré-processamento de dados	97
5.3	Métodos de Avaliação	102
5.3.1	Aceitação da Tecnologia	102
5.3.2	Treinamento e teste	106
5.3.3	Seleção do Modelo	107
5.3.4	Parâmetros de desempenho	109
5.4	Considerações Parciais	114
6	RESULTADOS E DISCUSSÃO	115
6.1	Avaliação de aceitação da Tecnologia	115
6.1.1	Avaliação do instrumento de medida	115
6.1.2	Avaliação qualitativa	116
6.2	Avaliação dos algoritmos de classificação	119
6.3	Avaliação do algoritmo de Predição	128
6.4	Discussão	131
6.4.1	Aceitação da Tecnologia	132
6.4.2	Experimentos de Classificação	134
6.4.3	Experimentos de Predição	138
6.5	Limitações	140
6.6	Considerações Parciais	142
7	CONCLUSÕES	143
7.1	Contribuições	144
7.2	Trabalhos Futuros	146
	REFERÊNCIAS	149

1 INTRODUÇÃO

Atualmente, *Diabetes mellitus* (DM) destaca-se como uma causa significativa de morbidade e mortalidade. Estimativas mundiais indicam que mais de 540 milhões de pessoas vivem com DM, e que poderia chegar a mais de 780 milhões em 2045 (SUN et al., 2022). Ainda assim, acredita-se que 50% das pessoas com diabetes não sabem que têm a doença (OGURTSOVA et al., 2022). Estima-se que 5,1 milhões de pessoas com idade entre 20 e 79 anos morreram de diabetes em 2013, podendo saltar da nona para a sétima maior causa de morte em todo o mundo até 2030 (OGURTSOVA et al., 2017; OGURTSOVA et al., 2022). No Brasil, a doença representa um problema de saúde significativo, ocupando em 2021, a sexta posição entre os países com maior número de diabéticos, com 15,7 milhões de casos entre indivíduos adultos e com previsão de alcançar 23,2 milhões no ano de 2045 (SUN et al., 2022; OGURTSOVA et al., 2017). O aumento da prevalência de obesidade, estilo de vida sedentário e o envelhecimento da população são considerados os principais fatores pelo aumento da incidência de DM em todo o mundo (DAMTIE et al., 2023; SZWARCOWALD; STOPA; MALTA, 2022). Esse cenário gera um alto custo social e financeiro para o paciente e para o sistema de saúde em geral. Estima-se que o DM foi responsável por 12% do total de internações não gravídicas e até 15,4% dos custos hospitalares do Sistema Único de Saúde (SUS) no período de 2008 a 2019 (BAHIA et al., 2023; FLORÊNCIO et al., 2021).

Nessa direção, com o objetivo de diminuir a sobrecarga do sistema de saúde, está sendo considerado um modelo orientado para o paciente, onde estes estão sendo equipados com conhecimento e tecnologias para desempenhar um papel mais ativo em seu monitoramento da saúde. Esse novo modelo promove o envolvimento ativo e o controle do paciente sobre sua própria saúde, por meio de educação e autogerenciamento adequados (PINTO et al., 2022; SNOWDON et al., 2022; MOONIAN et al., 2020). O objetivo da intervenção de autogerenciamento é apoiar o paciente a fazer as mudanças comportamentais necessárias para alcançar um controle ideal do diabetes, abrangendo os princípios de proatividade, independência, acessibilidade e custo-efetividade e usando uma ampla gama de tecnologias móveis, como smartphones, tablets e sensores vestíveis para o monitoramento contínuo do comportamento e dos sinais vitais (SOUZA-PEREIRA et al., 2020; ROSSETTI; TIASE, 2022).

O uso da tecnologia de Internet das Coisas da Saúde (IoHT) aparece como solução para impulsionar os serviços de saúde por meio de dispositivos portáteis, dispositivos de mão e smartphones (COSTA et al., 2018). Partindo de um sistema convencional baseado em hub para sistemas de saúde mais personalizados (PAGIATAKIS et al., 2020). No entanto, a transformação da tecnologia IoHT avançada em sistemas personalizados ainda é um desafio significativo na área. Alguns problemas incluem a falta de sensores médicos precisos e econômicos, arquiteturas de sistemas não padronizados de IoHT, heterogeneidade de dispositivos vestíveis conectados, multidimensionalidade dos dados gerados além uma alta demanda por interoperabilidade (BALOCH; SHAIKH; UNAR, 2018; QI et al., 2017; PASQUIER et al., 2018). Considerando

que o objetivo da computação onipresente em saúde é buscar a consciência do contexto (DEY; ABOWD; SALBER, 2001; TAVARES et al., 2016), alguns desafios se destacam, uma vez que a ideia de usar a consciência contextualizada em saúde é proporcionar inteligência e centrar os cuidados no paciente (MSHALI; LEMLOUMA; MAGONI, 2018; MONTORI et al., 2018). Como exemplos de desafios em relação ao uso da saúde sensível ao contexto, podemos mencionar a integração de dados de sensores heterogêneos e a eficiência energética de sensores (ESPOSITO et al., 2018; De La Iglesia et al., 2018; ZHANG et al., 2018), capacidade de resposta e robustez (MICHALAKIS; CARIDAKIS, 2017; RAHMANI et al., 2018; DEEN, 2015). Troca de contexto e entrega contínua de dados sem lacunas (BALOCH; SHAIKH; UNAR, 2018; BOBEK; NALEPA, 2017) aparecem como importantes dentro da construção do contexto onde o paciente sob cuidado está localizado (WETZELS et al., 2018; AMIN et al., 2018; KUIJS; ROSENCRANTZ; REICH, 2015; YI; NIE, 2018).

Por sua vez, na medicina, a Inteligência Artificial (IA) tem sido aplicado em diversos campos para diferentes propósitos. A IA é definida como "um campo da ciência e da engenharia preocupado com o entendimento computacional do que é comumente chamado de comportamento inteligente com a criação de artefatos que exibem esse comportamento"(JAVIDI et al., 2022; RIGLA et al., 2018; RAMESH et al., 2004). Métodos de IA em combinação com tecnologias mais recentes, incluindo a computação móvel e tecnologias de sensores, tem a capacidade de permitir a criação e prestação de serviços de gerenciamento de saúde melhores, para lidar com doenças crônicas (DRITSAS; TRIGKA, 2022; CONTRERAS; VEHI, 2018). A medicina moderna enfrenta desafios em relação a uma grande quantidade de aquisição de dados, análise e aplicação do conhecimento obtido na solução de problemas clínicos complexos, para esses fins, os recursos de IA são necessários (SARAVANAKUMAR; SABIBULLAH, 2022; ALDAHIRI; ALRASHED; HUSSAIN, 2021; FALLAH; KALHORI, 2017). A IA é composta de vários algoritmos e técnicas inteligentes, como aprendizado de máquina (ML), processamento de linguagem natural (PNL), lógica fuzzy (FL), sistemas especialistas (ES) e a mistura de dois ou mais métodos (multimetodos) (HAMET; TREMBLAY, 2017; NILSSON, 1998).

Métodos de ML comumente tem a capacidade de realizar a tarefas de previsão (quando a variável de resultado é um valor) ou de classificação (quando a variável de resultado é uma classe) (ABHARI et al., 2019). Estas tarefas podem ser aplicadas para previsão da probabilidade do surgimento da DM, triagem, diagnóstico, orientação para tratamento e gerenciamento de complicações. Vários métodos de ML foram aplicados para cumprir esses propósitos clínicos (YOUNUS et al., 2020; DWORZYNSKI et al., 2019; BRISIMI et al., 2019). Técnicas supervisionadas de classificação e clusterização foram aplicadas para identificar pacientes que poderiam ser hospitalizados devido a complicações atribuídas ao DM2 ou prever o risco de desenvolver comorbidades ligadas ao DM2 (BRISIMI et al., 2019; DWORZYNSKI et al., 2019; YOUNUS et al., 2020; MUNNA et al., 2020). Outros trabalhos realizam a previsão dos níveis de glicose baseados em LSTM (*long short-term memory*) e redes neurais recorrentes (RNNs)(FARUQUI et al., 2019), avaliar desempenho de programas de saúde multidisciplina-

res no controle da glicose usando árvores de regressão (YOU et al., 2019) ou ainda utilizando classificação e regressão logística para prever deterioração da saúde de pacientes com DM2 (RAMAZI et al., 2019). Embora os resultados dos algoritmos de classificação e predição tenham ficado aquém em comparação com outros estudos, nenhum trabalho havia desenvolvido a união de dados do RES de centro clínico nacional com monitoramento de pacientes para acompanhamento do curso da DM2. Ao mesmo tempo a pesquisa avalia a aceitação da tecnologia junto aos *stakeholders*, incluindo profissionais de saúde e pacientes, verificando os resultados de forma qualitativa e quantitativa, com resultados promissores com relação a adoção e aceitação da tecnologia.

1.1 Motivação

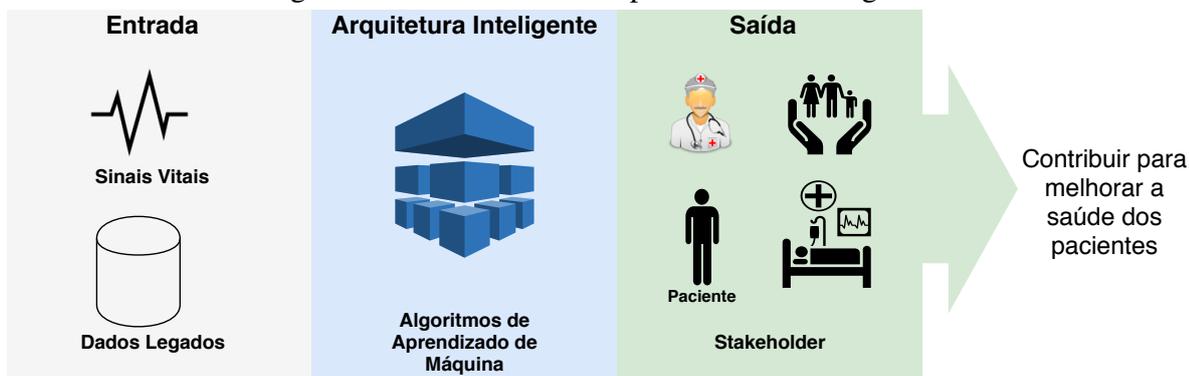
A digitalização do prontuário eletrônico pode ser colocada como uma revolução considerável da tecnologia da informação na medicina. A adoção generalizada de registros eletrônicos de saúde (RES) gerou conjuntos de dados massivos (LI et al., 2023; LIANG et al., 2021; MURDOCH; DETSKY, 2013). Devido ao crescente aumento da disponibilidade de dados, atingiu-se o limite da análise destes por humanos, exigindo uma dependência crescente de máquinas (TOPOL, 2019). O processo de diagnóstico e tratamento ainda estão concentrados na decisão do médico, e isso tende a permanecer inalterado. Porém, com os recentes avanços na capacidade de processamento computacional, armazenamento e o aperfeiçoamento de técnicas de processamento, algoritmos de inteligência artificial podem ser usados para executar tarefas tipicamente associadas à inteligência humana, com maior poder de processamento (VERMANI; PURVA, 2022; GOODFELLOW; BENGIO; COURVILLE, 2016).

Da mesma forma, com o aumento do uso de sensores vestíveis, que podem monitorar continuamente os sinais vitais - incluindo pressão arterial, frequência cardíaca, saturação de oxigênio, frequência respiratória e temperatura -, existe o potencial de antecipar um grande número de pacientes hospitalizados no futuro (ROSSETTI; TIASE, 2022; KAVAKIOTIS et al., 2017). O uso de dispositivos wearables e smartphones aparece como solução para impulsionar os serviços de saúde, partindo de um sistema convencional de saúde baseado em hub para sistemas de saúde mais personalizados. No entanto, a transformação da tecnologia de sensores avançada em sistemas personalizados ainda é um desafio significativo na área de saúde (ABDOLKHANI et al., 2023; SOUZA-PEREIRA et al., 2020; GUBERT; COSTA; RIGHI, 2020). O uso de técnicas de IA aplicadas na medicina e nos tratamentos e evolução de doenças crônicas (como a diabetes tipo 2) é uma oportunidade promissora, pois pode gerar conhecimento para apoiar a tomada de decisões dos stakeholder, como a equipe de atenção primária, médicos e hospitais.

Considerando os dados de RES e os dados gerados a partir do uso de sensores, podemos realizar a aplicação das abordagens de IA como forma de extrair padrões e criar modelos a partir destes dados. Isso é particularmente significativo quando o conjunto de dados pode atingir terabytes ou petabytes de volume (OIKONOMOU; KHERA, 2023; CONTRERAS; VEHI,

2018). Essa quantidade de dados fortaleceu consideravelmente a pesquisa orientada a dados em doenças crônicas. Como forma de contextualizar, propomos na Figura 1 um diagrama genérico de de fluxo de dados em um sistema inteligente de monitoramento para doenças crônicas não transmissíveis.

Figura 1 – Sistema de Acompanhamento Inteligente



Fonte: Elaborado pelo autor

Neste exemplo, as entradas são dadas pelo monitoramento de sinais vitais do paciente e por dados oriundos do registro eletrônico de saúde (RES) que, além de sinais vitais, pode contemplar dados como exames laboratoriais, uso de medicações, entre outros. Uma das aplicações mais importantes é o prognóstico e diagnóstico relacionados a doenças crônicas, que reduzem a qualidade de vida. Dentro desse espectro, está a diabetes mellitus tipo 2 (DM2) (CONTRE-RAS; VEHI, 2018; KAVAKIOTIS et al., 2017). O impacto social que a DM2 tem, faz dela um alvo de pesquisas. Aplicar métodos de aprendizado de máquina e de mineração de dados na pesquisa de DM2 é uma abordagem fundamental para extração de conhecimento (KAVA-KIOTIS et al., 2017; RIGLA et al., 2018). Para fins deste trabalho, a sigla DM2 se refere ao código da Organização Mundial da Saúde (OMS) para classificação estatística internacional de doenças e problemas relacionados à saúde CID10-E11¹, referente a Diabetes Mellitus Não-insulino-dependente (WELLS et al., 2011).

1.2 Problema de pesquisa

O diabetes mellitus tipo 2 (DM2) corresponde de 90 a 95% de todos os casos de DM. As escolhas terapêuticas para DM2 variam de acordo com as características individuais e clínicas dos pacientes (DIABETES, 2019). No Brasil o plano de tratamento baseia-se primariamente no

¹E11.0 DM não-insulino-dependente – com coma; E11.1 DM não-insulino-dependente – com cetoacidose; E11.2 DM não-insulino-dependente – com complicações renais; E11.3 DM não-insulino-dependente – com complicações oftálmicas; E11.4 DM não-insulino-dependente – com complicações neurológicas; E11.5 DM não-insulino-dependente – com complicações circulatórias periféricas; E11.6 DM não-insulino-dependente – com outras complicações especificadas; E11.7 DM não-insulino-dependente – com complicações múltiplas; E11.8 DM não-insulino-dependente – com complicações não especificadas; E11.9 DM não-insulino-dependente – sem complicações.

controle glicêmico e almeja prevenir complicações nos órgãos-alvo. Apesar da disponibilidade de diversas opções de tratamento, espera-se que a maioria dos pacientes desenvolva complicações microvasculares e macrovasculares ao longo do tempo (DIABETES, 2019).

Justamente por sua natureza crônica, que pode levar a complicações graves, o DM2 é uma condição onerosa tanto para os pacientes quanto para o sistema de saúde (GALLARDO-RINCÓN et al., 2021; BARCELO et al., 2017). Estudos avaliam o custo da DM2 entre pacientes ambulatoriais brasileiros e demonstram que estes aumentaram conforme a progressão do diabetes e a presença de complicações crônicas (PEREDA et al., 2022; FERRARI et al., 2022; BAHIA et al., 2011). O gerenciamento de uma doença crônica como o DM2, exige embasamento científico aliado a habilidades técnicas, prática clínica e capacidade de tomada de decisões. Está bem documentado que a tecnologia facilita a coleta, o processamento, o armazenamento e a troca de informações; englobando aplicativos com funções específicas em ambientes de saúde. Estudos identificam que os serviços de saúde que adotam a tecnologia de informação melhoram o monitoramento dos pacientes e a sua adesão ao tratamento, bem como reduzem as visitas às unidades de saúde. (DIABETES, 2019).

A tecnologia como fator de apoio ao tratamento e acompanhamento de pacientes DM2 tem sido foco de muitos trabalhos na área de saúde. Entretanto, poucos estudos visam explorar a opinião de médicos e pacientes DM2 sobre o uso da tecnologia em um mesmo experimento. Da mesma forma, os estudos que fazem uso de dados de pacientes, em geral, se utilizam de *datasets* disponibilizados com dados de pacientes de fora do Brasil, a exemplo do PIMA Indian (SMITH et al., 1988) e MIMMIC (JOHNSON et al., 2016), o que pode não refletir a realidade brasileira. Neste sentido, entendemos que o território faz diferença na percepção da avaliação das ferramentas tecnológicas e nos resultados de algoritmos.

De acordo com as considerações feitas acima, este estudo foi desenvolvido em parceria com um centro clínico brasileiro, onde são atendidos pacientes portadores da Diabetes Mellitus tipo 2. Em um primeiro momento, foi realizada uma avaliação da aceitação da tecnologia, por parte dos profissionais de saúde que trabalham na atenção primária, os quais realizam o acompanhamento dos pacientes com DM2. Após foram coletados dados históricos dos pacientes no RES disponibilizado pelo centro clínico, com os quais foram analisados através de algoritmos de aprendizado de máquina com o objetivo de classificar e prever o declínio de saúde dos pacientes. Da mesma forma, foi disponibilizado a um número de pacientes um aplicativo para coleta de sinais vitais no contexto do paciente, com o objetivo de complementar a série histórica de dados do paciente e melhorar resultados de predição de declínio da saúde. Por fim, foram realizadas entrevistas com pacientes que utilizaram o aplicativo, para verificar sua aceitação e opiniões sobre usabilidade.

O problema central da pesquisa desta proposta de tese é aplicar técnicas de aprendizado de máquina para o desenvolvimento de um sistema individualizado de acompanhamento e gerenciamento do desenvolvimento de comorbidades relacionadas ao DM2, considerando o monitoramento do paciente crônico com DM2 em seu contexto. A seguinte questão de pesquisa é

proposta:

Que estrutura deve ter uma arquitetura computacional, baseada em aprendizado de máquina, para detecção precoce de comorbidades decorrentes no tratamento da DM2?

Esta tese busca responder quais características são necessárias para que uma arquitetura computacional possa ser inteligente na detecção e controle de comorbidades ligadas ao DM2. A arquitetura proposta utiliza dois módulos que aplicam técnicas de aprendizado de máquina, para detectar crises em pacientes de DM2 e realizar a predição de futuras crises. Da mesma forma, a arquitetura envolve os diversos *stakeholders* no processo, como fornecedores de dados atualizados do paciente ou como consumidores das informações resultantes da arquitetura. Por fim a arquitetura mantém o modelo inteligente atualizado através da aplicação de aprendizado de máquina online, com dados coletados por sensores, nos diversos contextos do paciente. Sendo assim, a questão de pesquisa pode ser dividida em duas questões mais específicas:

Quais são as percepções dos profissionais de saúde, sobre o uso de tecnologia para acompanhamento de pacientes DM2?

A questão acima direciona para a aceitação e utilização da tecnologia no contexto do paciente DM2. Neste sentido, é importante fazer o levantamento das percepções dos *stakeholders* com participação no acompanhamento do curso da doença.

Quais são os melhores desempenhos de modelos ajustados para acompanhamento do paciente DM2?

A questão acima direciona a pesquisa para as definições técnicas a respeito da implementação da arquitetura. Quais técnicas de aprendizado de máquina foram utilizadas na implementação dos módulos inteligentes e como ocorre a integração dos módulos de aprendizado de máquina para implementação da arquitetura.

A hipótese para melhorar o acompanhamento do paciente, explorada neste estudo, é a predição precoce da deterioração da saúde dos pacientes. Para tornar essa previsão possível, desenvolvemos um conjunto de dados extraídos de um centro clínico de saúde (de pacientes portadores de DM2) com dados como sinais vitais, informações demográficas, adições, medicamentos utilizados e exames laboratoriais comuns. Posteriormente, submetemos esse conjunto de dados ao processamento com algoritmos de aprendizado de máquina mais utilizados em dados de saúde, segundo a literatura.

Foram empregados cinco métodos para abordar as questões de pesquisa específicas, mencionadas acima:

- Conduzir uma Revisão Sistemática da Literatura (RSL) para obter o estado da arte do objeto de estudo, identificar possíveis lacunas, desafios e tecnologias utilizadas;

- Propor uma arquitetura inteligente para auxiliar no autocuidado assistido em DM2, baseada nas lacunas identificadas;
- Utilização dos estudos clínicos para melhor compreender o comportamento e as percepções de médicos e outros profissionais de saúde com relação ao uso de tecnologia para acompanhamento do curso da doença;
- Avaliar a arquitetura com dados de pacientes de centro clínico brasileiro;
- Implementar um protótipo de coleta de dados no contexto do paciente, e testar em pacientes reais;

1.3 Organização do documento

O restante da proposta de tese está estruturado da seguinte forma. O capítulo 2 apresenta os conceitos fundamentais para entender os assuntos abordados nesta pesquisa. Temas relacionados ao Diabetes mellitus tipo 2 e sua abordagem pela atenção primária em saúde, com relação ao seu tratamento e sinais vitais e laboratoriais utilizados para o monitoramento. Por fim, são abordadas as técnicas de inteligência artificial e aprendizagem de máquina, que são utilizados.

O Capítulo 3 apresenta pesquisas específicas ao projeto proposto, detalhando o estado da arte da literatura sobre ciência de contexto aplicado a saúde. As abordagens e principais resultados, incluindo uma taxonomia para classificação da área, desafios e cenários relacionados, além da conexão entre a ciência de contexto e a DM2. O capítulo 4 detalha a configuração experimental da arquitetura AIDA, foco deste trabalho. com o detalhamento de seus componentes.

O capítulo 5 apresenta os materiais e a metodologia utilizada para avaliação da arquitetura AIDA. Também especifica os recursos iniciais que estão disponíveis, a estrutura utilizada e os casos de uso que consideramos durante o processo de desenvolvimento da arquitetura AIDA. O Capítulo 6 traz uma avaliação dos resultados dos testes realizados com os usuários (pacientes e profissionais de saúde) e dos algoritmos de aprendizado de máquina utilizados, analisando aspectos do estudo e comparando com outras pesquisas já realizadas na área. Por fim, o Capítulo 7 apresenta as conclusões do trabalho e sugere novas direções para este estudo, no futuro.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos essenciais usados na proposta de tese, que direcionam o desenvolvimento da arquitetura inteligente de auxílio no auto-cuidado e detecção do declínio da saúde de pacientes com diabetes tipo 2 (DM2). Uma caracterização da DM2 e os conceitos por trás da doença são detalhados na seção 2.1. Na seção 2.2 são apresentadas as técnicas de inteligência artificial, com ênfase no aprendizado de máquina utilizadas para o processamento de sinais vitais. Por fim, na Seção 2.3 são apresentadas as considerações parciais sobre o uso de técnicas de inteligência artificial aplicados em DM2.

2.1 Diabetes mellitus tipo 2

A diabetes mellitus tipo 2 (DM2) faz parte do grupo de Doenças Crônica não Transmissíveis (DCNTs). São doenças determinadas por diversos fatores de cunho sociais ou individuais, que se desenvolvem no decorrer da vida e são de longa duração. As principais DCNTs são as doenças cardiovasculares, doenças respiratórias crônicas, neoplasias e a diabetes mellitus. As DCNTs constituem o maior problema global de saúde e têm gerado elevado número de mortes prematuras, perda de qualidade de vida, com alto grau de limitação e incapacidade, além de serem responsáveis por impactos econômicos para famílias e comunidades, e para a sociedade em geral (MALTA et al., 2014).

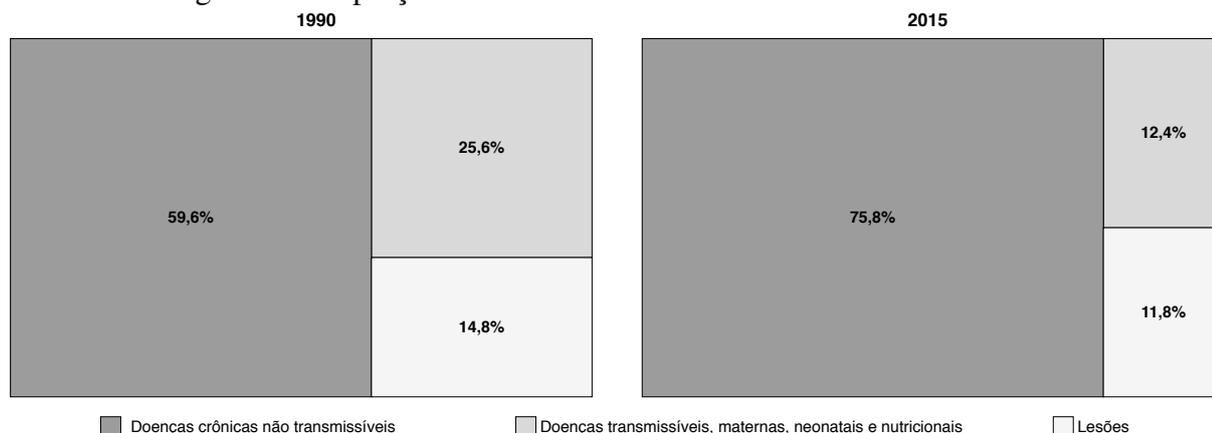
Atender às necessidades de pacientes com doenças ou comprometimentos crônicos é o maior desafio enfrentado pela prática médica organizada. Pesquisas mostram que proporções consideráveis de pacientes com doenças crônicas não estão recebendo terapia eficaz e têm pouco controle da doença. Resultados destes mesmos estudos mostram que programas eficazes de gerenciamento de doenças podem alcançar resultados substancialmente melhores do que os cuidados de intervenção e controle, que são os mais usuais (WAGNER, 1998).

Conforme pode ser visto na Figura 2, a proporção de mortes com origem em doenças crônicas aumentou significativamente entre os anos de 1990 e 2015. Com o objetivo de melhorar os cuidados, o Brasil apresentou o "Plano de Ações Estratégicas para o Enfrentamento das DCNT no Brasil 2011-2022"(MALTA; Morais Neto; Silva Junior, 2011). Dentre as metas propostas para o plano de enfrentamento das DCNTs podemos citar a redução da taxa de mortalidade prematura (<70 anos) por DCNT em 2% ao ano.

O Plano para enfrentamento das DCNTs aborda os quatro principais grupos de doenças (circulatórias, cânceres, respiratórias crônicas e diabetes) e seus fatores de risco em comum modificáveis (tabagismo, álcool, inatividade física, alimentação não saudável e obesidade) e define diretrizes e ações em três eixos: a) vigilância, informação, avaliação e monitoramento; b) promoção da saúde; c) cuidado integral (BRASIL. Ministério da Saúde, 2018; Inês Schmidt; Bartholow Duncan, 2011).

Com dados do ano de 2015, podemos classificar as DCNTs como tendo sido responsáveis

Figura 2 – Proporções de números de óbitos de acordo com as causas.



Fonte: Adaptado de (MALTA et al., 2017)

por 75,8% do total de óbitos na população no Brasil (SAÚDE, 2019). Neste sentido, segundo a Organização Mundial da Saúde (OMS), a diabetes aparece como responsável por 6% dos totais de óbitos (World Health Organization, 2016), conforme pode ser visualizado na Figura 3. Em termos de números de óbitos causados pelas DCNTs a diabetes melittus tipo 2 (DM2) aparece como a quarta maior causa de mortalidade, em números absolutos e proporcionalmente.

Com relação ao eixo do cuidado com o paciente portador de DCNT, é extremamente complexo, pois envolve um processo saúde-doença longo, exige mudanças no estilo de vida, necessidade de monitoramento e tratamento contínuo além de um amplo suporte dos serviços de saúde. É necessário a organização e estruturação dos serviços de saúde com conseqüente preparo dos profissionais que atendem visando uma adesão maior ao tratamento, a autonomia do paciente e melhora na qualidade de vida (Inês Schmidt; Bartholow Duncan, 2011)

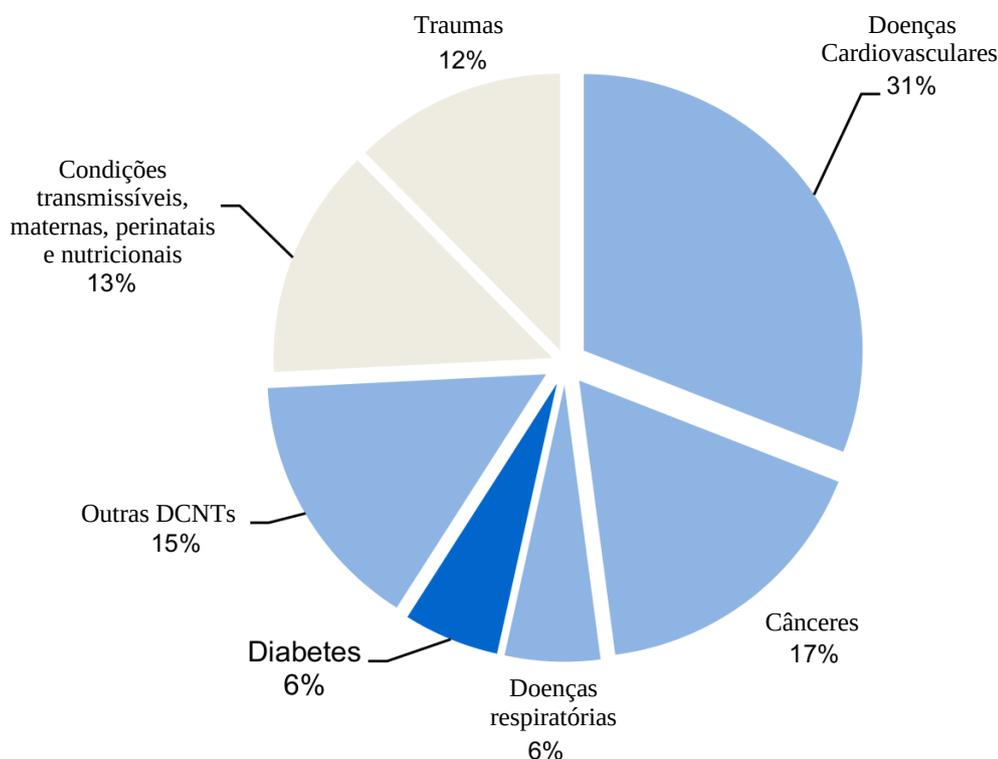
Da mesma forma, devido a natureza crônica da DM2, que pode levar a complicações graves, os cuidados de saúde tanto do paciente quanto do sistema de saúde se tornam uma condição onerosa. O custo do DM2 entre pacientes ambulatoriais brasileiros aumentam conforme a progressão do diabetes e a presença de complicações crônicas (SARAIVA et al., 2016).

As complicações crônicas incluem a nefropatia, com possível evolução para insuficiência renal; a retinopatia, com possibilidade de cegueira; e a neuropatia, com risco de úlceras nos pés, amputações e manifestações de disfunção autonômica. Além disso, o DM2 e a resistência à insulina encontram-se frequentemente associados com a hipertensão e as dislipidemias¹, aumentando a incidência da síndrome metabólica, os riscos para aterosclerose e doenças coronarianas (SAÚDE, 2019).

O tratamento de condições crônicas, por sua natureza, é diferente do tratamento dispensado a problemas agudos. Os sistemas de saúde atuais foram projetados para tratar dos problemas agudos e das necessidades prementes dos pacientes (diagnóstico - atenuação de sintomas - cura), o que não se encaixa quando falamos de doenças crônicas como o diabetes (JORDAN, 2003).

¹**Dislipidemia** é elevação de colesterol e triglicérides no plasma ou a diminuição dos níveis de HDL que contribuem para a aterosclerose.

Figura 3 – Proporção de mortes causadas pela DM2 no Brasil



Fonte: Adaptado de (World Health Organization, 2016)

2.1.1 Tratamento da diabetes mellitus tipo 2

Atender as necessidades de pacientes com doenças ou comprometimentos crônicos é o maior desafio enfrentado pela prática médica organizada, já que cuidado usual não está fazendo o trabalho. Dezenas de pesquisas e auditorias revelaram que proporções consideráveis de pacientes com doenças crônicas não estão recebendo terapia eficaz, têm pouco controle da doença e estão descontentes com seus cuidados (WAGNER, 1998; ALIZADEH et al., 2023; NOORDMAN et al., 2024).

Segundo (WAGNER, 1998) a melhoria real dos resultados ocorrerá somente quando os sistemas clínicos se reconfigurarem especificamente para atender às necessidades e preocupações dos pacientes com doenças crônicas. Nas doenças crônicas, o sucesso do tratamento depende fortemente da participação e do envolvimento do usuário enquanto sujeito ativo de seu tratamento (Brasil, Ministério da Saúde, 2013; PRAHALAD et al., 2024).

Uma atitude de autocuidado que direcione a um estilo de vida mais saudável e uma adesão ao tratamento não depende apenas de uma prescrição profissional, mas de uma conscientização do usuário sobre a sua condição de saúde. O autocuidado apoiado enseja uma colaboração estreita entre os usuários e a equipe de saúde, com a definição de vários pontos, como o estabelecimento de metas, o monitoramento, a instituição de planos de cuidados e a resolução de problemas ao longo do processo de manejo (Brasil, Ministério da Saúde, 2013).

Na Figura 4 pode ser visto um modelo que demonstra as necessidades de cuidados de pacientes com doenças crônicas, sugerindo que as interações paciente-provedor possam melhorar a partir dos seguintes requisitos:

- Organização da atenção à saúde: ter processos e incentivos bem desenvolvidos para fazer alterações no sistema de prestação de cuidados;
- Autocuidado apoiado: garantir apoio de autogestão comportamental sofisticado, priorizando o aumento da confiança e das habilidades dos pacientes, para que eles possam ser o gerente final de sua doença;
- Desenho da linha de cuidado: reorganizar os sistemas de consultas e acompanhamento, para atender às necessidades de pacientes com doenças crônicas;
- Suporte às decisões clínicas: desenvolver e implementar diretrizes baseadas em evidências e melhorar a interação entre generalistas e especialistas;
- Sistema de informações clínicas: aprimorar os sistemas de informação para facilitar o desenvolvimento de registros de doenças, sistemas de rastreamento e fornecer feedback sobre o desempenho.

Os resultados funcionais e clínicos melhores, propostos pelo modelo, passam pelo requisito de interações mais produtivas entre um paciente informado e ativo no seu tratamento e uma equipe de atenção proativa e preparada (WAGNER, 1998).

Do mesmo modo, o tratamento do diabetes mellitus tipo 2 necessita de um acompanhamento mais detalhado por parte do sistema de saúde (GUIDONI et al., 2009). A maioria dos portadores do DM2 apresenta comorbidades, tais como: obesidade, HAS² e dislipidemia. Portanto, o cuidado requer equipe multiprofissional, com abordagem interdisciplinar, colaborativa e integrada, estimulando um papel ativo do paciente no cuidado de sua patologia (ARAÚJO et al., 1999).

Conforme pode ser visto na Figura 5, fluxograma do tratamento de pacientes com diabetes tipo 2, o tratamento do paciente já diagnosticado com DM2 consiste em estágios de interações do paciente e os profissionais de saúde. O primeiro passo terapêutico para mais de 80% dos pacientes recém-diagnosticados se baseia em medidas para controle de peso, dieta hipocalórica e aumento de atividade física. Se o paciente não alcançar a meta glicêmica em até 3 meses, com as mudanças no estilo de vida, ocorre a primeira intervenção com medicamento, além de continuar com as medidas não farmacológicas. As interações cíclicas continuam ocorrendo a cada 3 ou 6 meses com a adição de outros fármacos e manutenção das mudanças de estilo de vida. Nos intervalos das interações, caso as metas sejam atingidas, ocorre o monitoramento do paciente (DUNCAM et al., 2013)

²Hipertensão Arterial Sistêmica

Figura 4 – Modelo para melhoria do tratamento de doenças crônicas



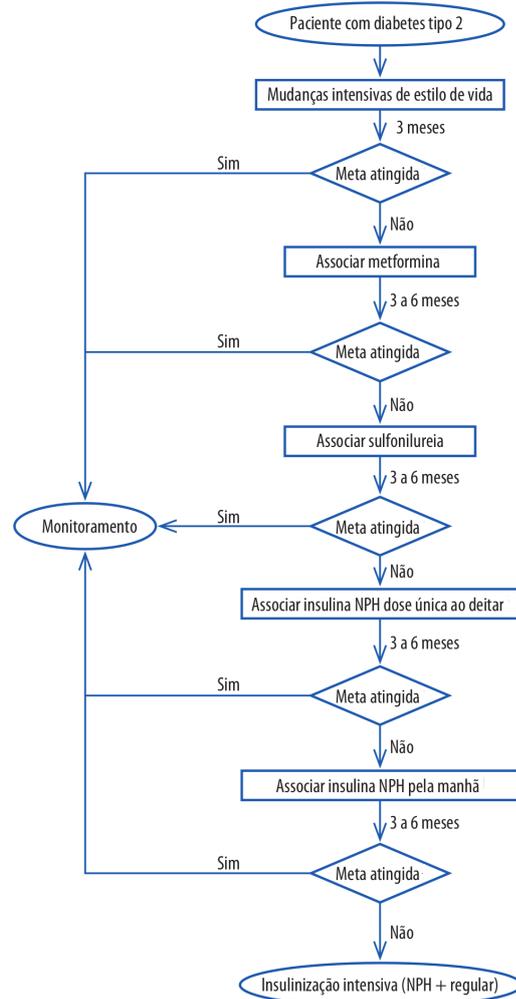
Fonte: Adaptado de (WAGNER, 1998)

As complicações agudas e crônicas do diabetes causam alta morbimortalidade o que acarreta altos custos para os sistemas de saúde. Gastos relacionados ao diabetes mundialmente, em 2010, foram estimados em 11,6% do total dos gastos com atenção em saúde (Internation Diabetes Federation, 2019). Dados brasileiros sugerem valores semelhantes (ROSA et al., 2014; Internation Diabetes Federation, 2019). Estudo realizado pela OMS mostrou que os custos governamentais de atenção ao DM variam de 2,5% a 15% dos orçamentos anuais de Saúde, e os custos podem exceder, em até cinco vezes, os custos diretos de atenção à saúde (World Health Organization, 2016).

Estudos internacionais sugerem que o custo dos cuidados relacionados ao diabetes é cerca de duas a três vezes superior aos dispensados a pacientes não diabéticos e está diretamente relacionado com a ocorrência de complicações crônicas (Internation Diabetes Federation, 2019).

A análise do impacto econômico e social do número crescente de pessoas que vivem com DM2 mostra a necessidade da implementação e de pesquisas em áreas que minimizem as dificuldades dessas pessoas e de suas famílias em gerenciar a doença, e propiciem a manutenção da

Figura 5 – Fluxograma de tratamento de pacientes com DM2



Fonte: Adaptado de (DUNCAM et al., 2013)

sua qualidade de vida. Da mesma forma, o sistema público de saúde apresenta a necessidade de minimizar o impacto da DM2 nos seus custos e agilizar o acompanhamento destes pacientes.

Em (GUIDONI et al., 2009) são citados quatro pontos fundamentais para redução dos custos e um melhor cuidado do DM2, o rastreamento e diagnóstico precoce, a garantia de acesso e utilização do serviço de saúde e a qualidade do cuidado prestado. Existe uma evolução natural da enfermidade, que pode ser acelerado quando o esquema terapêutico não é cumprido adequadamente. Em geral a adesão ao tratamento é em torno de 50% dos pacientes com diagnóstico positivo (DUARTE-RAMOS; CABRITA, 2006). Dentro desta perspectiva a tecnologia aparece como uma ferramenta útil para aumentar a adesão, otimizar o monitoramento dos pacientes durante e entre as interações com a equipe de saúde e, por fim, reunir dados mais precisos para o gerenciamento dos custos de atenção primária à saúde.

2.1.2 Sinais vitais e laboratoriais para monitoramento de DM2

Sinais vitais são indicadores das funções vitais e podem orientar o diagnóstico inicial e o acompanhamento da evolução da saúde do paciente. São mais significativos quando obtidos em série, possibilitando o acompanhamento de suas variações. As alterações das funções corporais geralmente se refletem na temperatura do corpo, na respiração e na pulsação arterial, podendo indicar enfermidades. Essas medidas fornecem informações muito importantes sobre as condições de saúde dos pacientes, pois é um método eficiente de monitoramento (GOLDMAN; SCHAFER, 2011).

Desta forma, os sinais vitais podem desempenhar uma função importante na tomada de decisões médicas. Por exemplo, uma temperatura corporal anormal baixa pode ser muito séria e até fatal. A baixa temperatura corporal pode ocorrer devido à exposição ao frio, estados de choque, ou em certos distúrbios metabólicos, como diabetes ou hipotireoidismo (SMITH; ROBERTS, 2011).

Os quatro principais sinais vitais são: a) pressão arterial (diastólica e sistólica), b) temperatura corporal, c) frequência cardíaca e d) frequência respiratória (BICKLEY, 2013). A pressão arterial (PA) consiste na força do sangue exercida nas paredes das artérias à medida que é bombeada para fora do coração, se divide em sistólica e diastólica. A primeira é a medida da pressão no pico da contração ventricular esquerda e a segunda é a pressão de repouso na parede arterial, ocorrendo durante o relaxamento ventricular. Para medir a pressão nos vasos sanguíneos, o medidor usa uma escala de milímetros de mercúrio (mmHg) (SMITH; ROBERTS, 2011)

A temperatura corporal é regulada pelo centro termorregulador do hipotálamo que equilibra a produção de calor e a perda de calor do corpo (SMITH; ROBERTS, 2011). Existem vários métodos comuns para avaliação da temperatura corporal, incluindo oral, retal, artéria axilar, timpânica e temporal (BICKLEY, 2013). A frequência cardíaca ou frequência de pulso, reflete o número de vezes que o coração bate por minuto. O coração é um órgão muscular que se contrai e relaxa, forçando o sangue do coração para as artérias sistêmicas. O pulso é a força do sangue contra as paredes das artérias e gera uma onda rítmica de pressão que pode ser sentida e medida em vários pontos do corpo como pulso (BICKLEY, 2013).

A frequência respiratória é controlada na mesma porção do cérebro que controla a profundidade da respiração, sendo que os principais órgãos do sistema respiratório são os pulmões (BICKLEY, 2013). Uma respiração consiste nas fases de inspiração e expiração e, para um adulto saudável em repouso, essas fases são iguais em comprimento (GOLDMAN; SCHAFER, 2011).

Para cada sinal vital, existem faixas consideradas normais para uma parcela da população saudável, conforme pode ser visto na Tabela 1. No entanto é importante considerar que as faixas normais de sinais vitais mudam conforme a idade, sexo, peso e saúde em geral (GOLDMAN; SCHAFER, 2011). O uso de faixas normais estabelecidas para cada sinal vital pode auxiliar médicos e profissionais de saúde a descobrir alguma anormalidade. Quando medidos periodi-

camente, os sinais vitais são uma ferramenta de monitoramento que pode significar a progressão de um processo de doença e a eficácia do tratamento e acompanhamento (SMITH; ROBERTS, 2011).

Tabela 1 – Intervalos de referência para sinais vitais e laboratoriais em adultos diabéticos

Sinal Vital	Intervalo normal	Unidade de medida
P. Arterial (sistólica)	90-130	mmHg
P. Arterial (diastólica)	60-90	mmHg
Temperatura	37,2 - 37,7	°C (graus celsius)
Frequência Cardíaca	60 - 100	batimentos/minuto (bpm)
Oxigenação (SpO2)	95% - 99%	
Glicose	70 - 120	mg/dl

Fonte: Adaptado de (GOLDMAN; SCHAFER, 2011)

A Tabela 1 apresenta mais duas informações relevantes, que é a saturação de oxigênio (SpO2) e a glicose (MOWER et al., 1998; DEFRONZO et al., 2015). A saturação de oxigênio é uma medida da quantidade de hemoglobina ligada ao oxigênio em comparação com a quantidade de hemoglobina que permanece sem ligação. Existem duas medidas: a saturação arterial de oxigênio (SaO2) que é expressa como uma porcentagem geral de locais de ligação da hemoglobina que são ocupados por oxigênio (medida invasiva) e a saturação periférica de oxigênio (SpO2) que é um valor obtido do sangue capilar periférico usando a oximetria de pulso que geralmente reflete de perto os níveis que seriam encontrados no sangue arterial (não invasiva) (SMITH; ROBERTS, 2011). Um estudo que apresenta a influência das alterações da microcirculação nas trocas de fluídos e oxigênio em pacientes com DM2, pode ser visto em (TANG; HE, 2018).

A glicose é fundamental para manter a atividade metabólica e a função celular. A coleta de uma amostra de glicose no sangue fornecerá informações valiosas sobre um paciente que está gravemente doente ou sobre um paciente com diabetes (SMITH; ROBERTS, 2011). O monitoramento do açúcar no sangue de um paciente fornece uma indicação de como o corpo está controlando o metabolismo da glicose. Existem duas maneiras de obter uma amostra de sangue; o primeiro método é colher uma amostra de sangue da veia, que é então enviada ao laboratório para ser testada; ou um método mais simples e rápido é uma amostra capilar. Uma amostra capilar de sangue é obtida dos minúsculos vasos sanguíneos chamados capilares que estão próximos à superfície da pele do dedo. (DEFRONZO et al., 2015; SMITH; ROBERTS, 2011)

Além das informações listadas na Tabela 1, exames laboratoriais são extremamente relevantes e são usados como complemento à clínica médica, guiando a busca pelo diagnóstico e tratamento. Entre estes, podemos destacar o exame de hemoglobina glicada (HbA1c) e o exame qualitativo de urina (EQU). A HbA1c é composta por diversas frações de hemoglobina e é o melhor exame para controle glicêmico a longo prazo, sem sofrer influências de refeições re-

cente do paciente ou alterações glicêmicos agudas (GOLDMAN; SCHAFER, 2011). Já o EQU consiste em uma análise de amostra de urina, observando a cor da amostra, presença de células, proteínas, bactérias, entre outras substâncias, mostrando ser um exame capaz de conduzir diagnósticos diferenciais (MCPHERSON; PINCUS, 2017).

A albumina consiste em uma proteína sanguínea sintetizada no fígado e usada no transporte de substâncias apolares, ou seja, que não tem carga positiva ou negativa, como hormônios, ácidos graxos e fármacos. Os glomérulos renais são responsáveis pela filtração dos compostos sanguíneos, evitando que eles sejam excretados pela urina, porém, em situações de inflamação ou infecção glomerular, ocorre perda da albumina na urina - albuminúria (LIEBERMAN; PEET, 2018). A albuminúria é um importante marcador de função renal, detectado em doença renal, com alteração na taxa de filtração glomerular. Seus valores também estão relacionados à uma maior incidência de eventos cardiovasculares. Em pacientes com diabetes, a presença de albuminúria representa um aumento significativo nas probabilidades de desenvolver neuropatia (GOLDMAN; SCHAFER, 2011).

2.2 Inteligência artificial

Um ponto central da arquitetura AIDA é o uso de técnicas de inteligência artificial como suporte a tomada de decisões nos cuidados do paciente. Desta forma, esta seção irá explorar os conceitos relacionados à inteligência artificial, tipos de aprendizado, redes neurais, online learning, e a relação destes com os diversos *stakeholders* envolvidos no tratamento do paciente portador de diabetes tipo 2.

Podemos encontrar várias definições diferentes para inteligência artificial. Começamos pela definição primordial de Alan Turing como sendo a capacidade de obter desempenho no nível humano em tarefas cognitivas (TURING, 1950). Segundo (LUGER, 2013), a IA pode ser definida como um ramo da ciência da computação que trata da automação do comportamento inteligente. Por sua vez (HAYKIN, 2008) coloca que o campo da inteligência artificial tenta não apenas compreender, mas também construir entidades inteligentes. Independente da definição mais utilizada, o objetivo de um sistema inteligente é interpretar corretamente dados externos, aprender a partir desses dados e utilizar essas aprendizagens para atingir objetivos e tarefas específicos através de adaptação flexível (ALPAYDIN, 2016)

O aprendizado envolve generalizar a partir da experiência. O desempenho do sistema inteligente deve não somente melhorar na execução da mesma tarefa, mas também em tarefas semelhantes, no mesmo domínio (LUGER, 2013). Aprendizado de máquina é, dentro da área de inteligência artificial, um conjunto de métodos focados em como construir sistemas de computador que possam melhorar automaticamente com a experiência. Muitas das aplicações de aprendizado de máquina objetivam auxiliar em tarefas relacionadas ao processamento de linguagem natural, controle de robôs e visão computacional, dentre outros. Para estas aplicações, e outras de caráter prático, pode ser mais simples treinar um sistema, através do uso de exemplos

de comportamento desejado de entrada e saída, do que programá-lo manualmente, antecipando a saída desejada para todas as entradas possíveis (JORDAN; MITCHELL, 2015; MUELLER; GUIDO, 1997; ALPAYDIN, 2016). Desta forma, as próximas seções do texto focam em descrever os tipos de aprendizagem e suas implicações para a arquitetura propostas neste trabalho.

2.2.1 Tipos de aprendizado

Uma das formas de categorizar os algoritmos de aprendizado de máquina é pelo seu tipo de aprendizado. Neste sentido, podemos separar os algoritmos em três principais categorias, a aprendizagem supervisionada (algoritmos de classificação e previsão com base em exemplos anteriores), a não supervisionada (capacidade de encontrar padrões) e aprendizagem por reforço (uso de sequências de recompensas e punições com objetivo de formar uma estratégia de operação em um domínio específico). Qual tipo de algoritmo aplicar para a solução de um determinado problema, depende substancialmente de uma análise do problema e dos dados disponíveis, para que possa ser criado um ambiente de aprendizado adequado (ALPAYDIN, 2016; ALANAZI; ABDULLAH; QURESHI, 2017; RAMESH et al., 2004).

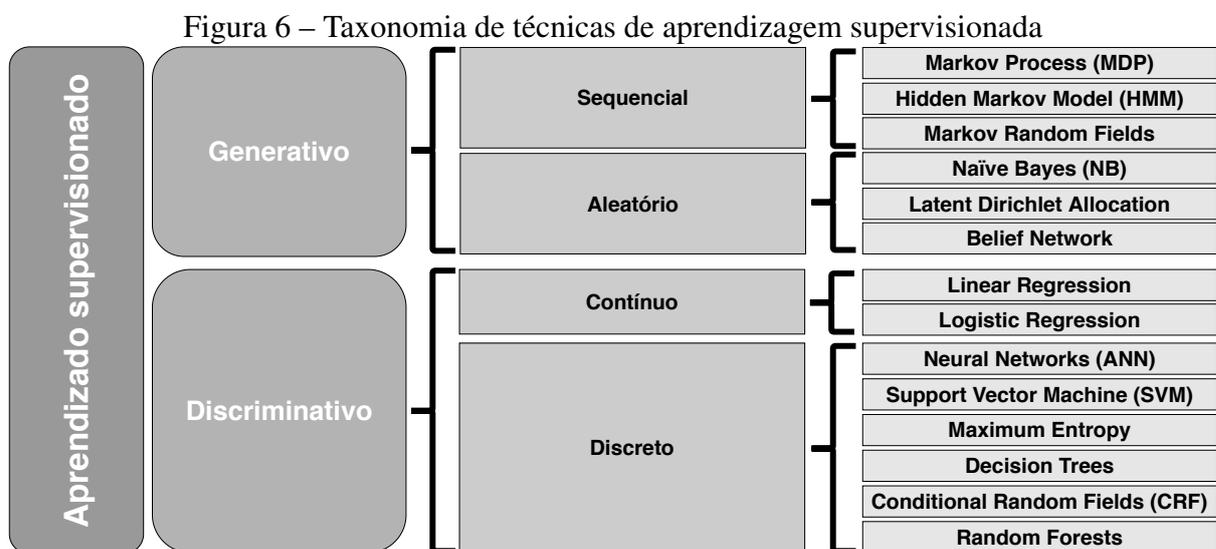
2.2.1.1 Supervisionado

O aprendizado de máquina supervisionado é um dos tipos de algoritmo de aprendizado de máquina mais utilizado. O aprendizado supervisionado é usado sempre que queremos prever um determinado resultado dado um conjunto de exemplos de entradas X e saídas Y . Construímos um modelo de aprendizado supervisionado a partir desses pares de X e Y , que compõem o conjunto de treinamento. O objetivo é fazer previsões precisas de dados novos e nunca antes vistos (ZHANG; Ali Babar, 2013). Geralmente o aprendizado supervisionado requer esforço humano para construir o conjunto de treinamento, mas depois automatiza e frequentemente acelera uma tarefa trabalhosa ou inviável. Em muitos casos, as saídas y devem ser fornecidas por um “supervisor” humano, pois podem ser difíceis de se coletar automaticamente. Mas mesmo quando as metas do conjunto de treinamento forem coletados automaticamente, o termo “supervisionado” ainda se aplica (GOODFELLOW; BENGIO; COURVILLE, 2016; MUELLER; GUIDO, 1997).

Em outras palavras, o aprendizado supervisionado é uma abordagem de aprendizado de máquina deduzida de dados de treinamento rotulados, compostos por um conjunto de exemplos de treinamento. Cada exemplo neste conjunto de dados compreende um par de objetivos de entrada, um vetor de entrada e um valor de saída. Neste tipo de aprendizado, um algoritmo analisa os dados de treinamento e produz uma função inferida chamada classificador. Essa função deduzida deve prever com precisão o valor de saída para qualquer objeto de entrada adequado. Esta abordagem significa que há uma generalização dos dados de treinamento para situações anteriormente não observadas (ALANAZI; ABDULLAH; QURESHI, 2017; KAVAKIOTIS et

al., 2017). Desta forma, após o treinamento o modelo pode ser salvo e utilizado para prever resultados com base em novos dados (GOODFELLOW; BENGIO; COURVILLE, 2016).

Algoritmos de aprendizado de máquina supervisionados podem ser divididos em duas categorias, modelos generativos e modelos discriminativos (NICOLAS, 2014). As diferenças básicas entre eles ocorrem com relação a alguns critérios como a acurácia, requisitos de modelagem, custo computacional e restrições. Uma taxonomia parcial de algoritmos de aprendizagem supervisionada pode ser vista na Figura 6, onde o modelo generativo é decomposto em aleatório e sequencial e o modelo discriminativo em processos contínuos (regressão) e processos discretos (classificação) (NICOLAS, 2017).



Fonte: Adaptado de (NICOLAS, 2014)

2.2.1.2 Não supervisionado

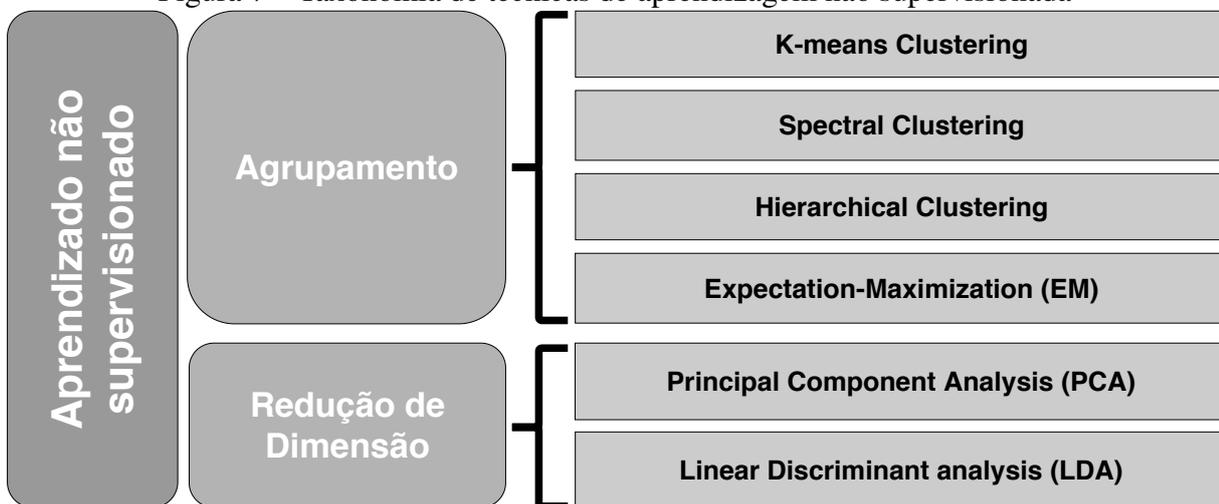
Diferentemente do aprendizagem supervisionada, a aprendizagem não supervisionada utiliza dados onde apenas a entrada x é conhecida. (ANDREAS; GUIDO, 2017). Na aprendizagem não supervisionada, o sistema tenta descobrir a estrutura oculta de dados ou associações entre variáveis. Nesse caso, os dados de treinamento consistem em instâncias sem rótulos correspondentes (KAVAKIOTIS et al., 2017). O objetivo da aprendizagem não supervisionada é descobrir padrões de regularidade e de irregularidades em um conjunto de observações. A aprendizagem não supervisionada está relacionada à estimativa de densidade no campo da estatística (NICOLAS, 2014).

Algoritmos não supervisionados podem ser categorizados em dois tipos de técnica, a clusterização e a redução de dimensão, conforme pode ser visto na taxonomia apresentada na Figura 7. A função de clustering é descobrir grupos de observações relacionadas nos dados de treinamento (HACKELING, 2014). Essa tarefa, denominada agrupamento ou análise de agrupamento, atribui observações a grupos de modo que as observações dentro dos grupos sejam

mais semelhantes entre si com base em alguma medida de similaridade do que as observações em outros grupos. O clustering é frequentemente usado para explorar um conjunto de dados. (ANDREAS; GUIDO, 2017)

A técnica de redução de dimensionalidade, outra categoria da aprendizagem não supervisionada, consiste em obter uma representação de alta dimensão dos dados, consistindo em muitos recursos, e encontrar uma nova maneira de representar esses dados que resumem as características essenciais, com menos recursos (ANDREAS; GUIDO, 2017). Em outras palavras, alguns problemas podem conter milhares ou mesmo milhões de variáveis explicativas, o que pode se tornar computacionalmente dispendioso. A redução de dimensionalidade é o processo de descobrir as variáveis explicativas que respondem pelas maiores mudanças na variável de resposta (HACKELING, 2014).

Figura 7 – Taxonomia de técnicas de aprendizagem não supervisionada



Fonte: Adaptado de (NICOLAS, 2014)

2.2.1.3 Aprendizagem por reforço

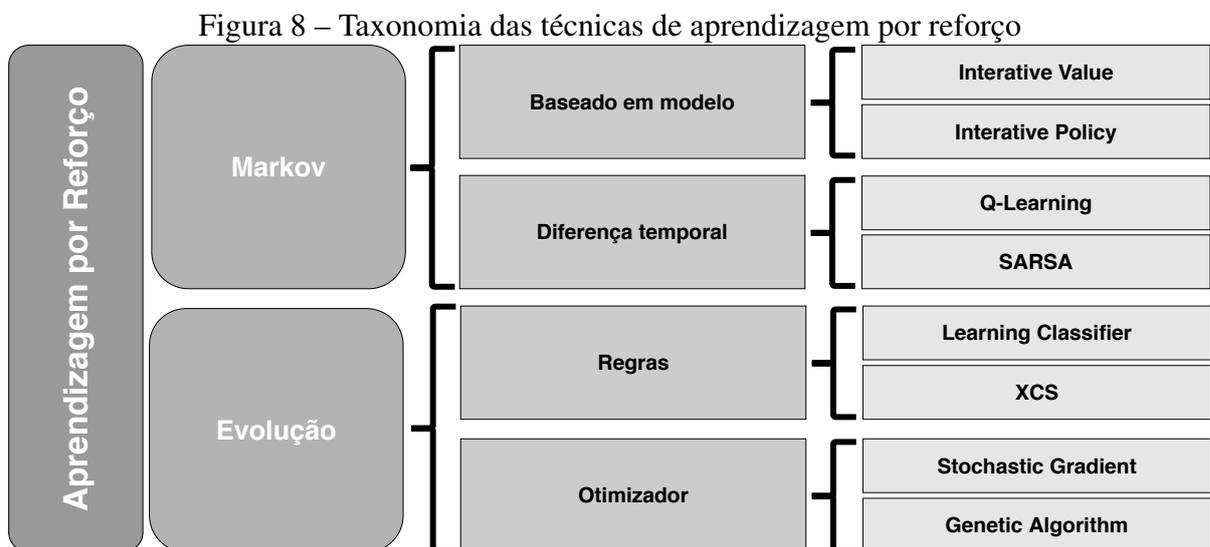
A abordagem de aprendizagem por reforço é mais focada no aprendizado direcionado a objetivos da interação frente a outras abordagens de aprendizado de máquina (SUTTON; BARTO, 2018). A aprendizagem por reforço não é tão bem entendido fora dos domínios da robótica ou da estratégia de jogos (NICOLAS, 2017). Para alguns tipos de aplicativos, os algoritmos de aprendizagem por reforço emitem um conjunto de ações recomendadas para execução do sistema adaptativo, que na forma mais simples, calcula ou estimam o melhor curso de ação (ALPAYDIN, 2016; NICOLAS, 2017; SUTTON; BARTO, 2018).

O aprendizado por reforço é aprender o que fazer, ou seja, como mapear situações para ações (SUTTON; BARTO, 2018). O objetivo é fazer com que o computador aprenda a fazer algo que não dizemos como fazer. A abordagem consiste em não fornecer categorizações explícitas, mas utilizar algum tipo de recompensa para indicar sucesso (ZHANG; Ali Babar, 2013; SUTTON;

BARTO, 2018). Esse tipo de treinamento geralmente se encaixa na estrutura do problema de decisão porque o objetivo não é produzir uma classificação, mas tomar decisões que maximizem as recompensas. Generalizando bem para problemas do mundo real, já que aplica recompensas por realizar determinadas ações e punições por outras (ZHANG; Ali Babar, 2013). Desta forma, aprendizagem por reforço é uma abordagem indicada para a implementação de agentes inteligentes, como dispositivos móveis autônomos (robôs).

Como exemplo, podemos pensar em um jogo de xadrez. Partindo do princípio que podemos observar todas as posições das peças do tabuleiro (nossas e do oponente), o objetivo é decidir as jogadas para ganhar o jogo. Por dois motivos, não é viável utilizar um aprendizado supervisionado. Primeiro porque é muito dispendioso ter supervisão durante todo o jogo, indicando a melhor jogada para cada estado do tabuleiro. Segundo, porque em muitos casos, a melhor jogada depende dos próximos movimentos (ALPAYDIN, 2016). Um único movimento não conta, uma sequência de movimentos é boa se, ao final delas, ganhamos o jogo. O único feedback real é no final, quando ganhamos ou perdemos o jogo. A idéia é descobrir quais ações podem afetar não apenas a recompensa imediata, mas também a próxima situação, e com isso todas as recompensas subsequentes (SUTTON; BARTO, 2018).

Da mesma forma que nas duas estratégias de aprendizado anteriores, os modelos de aprendizado por reforço podem ser categorizados, neste caso, como markovianos ou evolutivos (NICOLAS, 2014). Para melhor compreensão, mostramos na Figura 8 uma possível taxonomia para algoritmos de aprendizagem por reforço.



Fonte: Adaptado de (NICOLAS, 2014)

2.2.2 Objetivos na aprendizagem de máquina

O objetivo no aprendizado de máquina, é um tipo de previsão ou inferência que está sendo feita, com base no problema ou pergunta que está sendo colocada, e nos dados disponíveis.

Dentro do escopo do aprendizado de máquina, duas das tarefas realizadas são: a) classificação e b) regressão (HACKELING, 2014; ALPAYDIN, 2014). Na tarefa de classificação, a técnica necessita, com base em um conjunto de entradas, prever rótulos discretos para cada instância. Por sua vez, a regressão consiste prever rótulos contínuos para cada instância (ALANAZI; ABDULLAH; QURESHI, 2017; RAMESH et al., 2004). Um exemplo é uma coleção de recursos que foram medidos quantitativamente a partir de algum objeto ou evento que queremos que o sistema de aprendizado de máquina processe. Normalmente representamos um exemplo como um vetor $x \in \mathbb{R}^n$ onde cada entrada x_i do vetor é outro recurso. (GOODFELLOW; BENGIO; COURVILLE, 2016)

2.2.2.1 Classificação

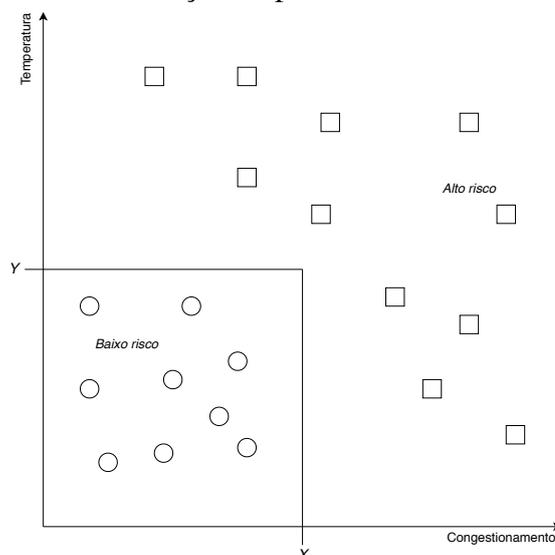
A classificação é uma sub-categoria de aprendizagem supervisionada. Classificação é o processo de tomar algum tipo de entrada e atribuir um rótulo a ela. Sistemas de classificação são usados geralmente quando as previsões são de natureza distinta, ou seja, um simples “sim ou não”. A classificação consiste em especificar a quais categorias de k algumas entradas pertencem. Para resolver esta tarefa, geralmente é solicitado ao algoritmo de aprendizado que produza uma função $f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$. Quando $y = f(x)$, o modelo atribui uma entrada descrita pelo vetor x a uma categoria identificada pelo código numérico y . Existem outras variantes da tarefa de classificação, por exemplo, em que f gera uma distribuição de probabilidade nas classes (GOODFELLOW; BENGIO; COURVILLE, 2016).

Um exemplo de objetivo da classificação pode ser extrair conhecimento de dado histórico. Pode ser construído um classificador para identificar uma doença a partir de um conjunto de sintomas. Com um conjunto de informações coletadas sobre dados fisiológicos, como temperatura corporal (variável contínua), congestionamento (variável discreta) e o diagnóstico final, é possível criar um modelo que afirme com uma determinada probabilidade que o paciente tem gripe. (NICOLAS, 2017).

Na figura 9, separamos os paciente em duas classes: paciente de baixo e alto risco de estar gripado. As informações sobre um paciente compõem a entrada para o classificador, cuja tarefa é atribuir a entrada a uma das duas classes. Após o treinamento com os dados históricos, uma regra de classificação apreendida pode ter a forma SE *temperatura* > Y E *congestionamento* > X ENTAO *alto – risco* SENAO *baixo – risco* para valores adequados de X e Y . Este é um exemplo de um discriminante, que é uma função que separa os exemplos de diferentes classes (ALPAYDIN, 2014).

Uma vez que o modelo é treinado utilizando observações históricas e validado contra observações históricas, ele pode ser usado para prever algum resultado. Um médico pode coletar sintomas de um paciente, como temperatura corporal e congestão nasal, e antecipar o estado de sua saúde (NICOLAS, 2017).

Figura 9 – Classificação de pacientes com alto e baixo risco de gripe



Fonte: Adaptado de (ALPAYDIN, 2016)

Outra forma de aplicar a classificação, ao invés de tomar uma decisão do tipo 0/1 (baixo risco/alto risco), podemos querer calcular uma probabilidade $P(Y|X)$, onde X são os atributos do paciente e Y é 0 ou 1, respectivamente, para baixo e alto risco. Nesta perspectiva, podemos ver a classificação como aprendizado de uma associação de X a Y . Então, para um dado $X = x$, se tivermos $P(Y = 1|X = x) = 0,8$, dizemos que o paciente tem uma probabilidade de 80% de alto risco ou uma probabilidade equivalente de 20% de baixo risco. Pode-se então utilizar esta informação para auxiliar no diagnóstico como positivo ou negativo para gripe. (ALPAYDIN, 2014)

2.2.2.2 Regressão

Nesse tipo de tarefa é solicitado ao programa de computador prever um valor numérico, com base em alguma entrada. Para resolver essa tarefa, é solicitado ao algoritmo de aprendizado gerar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Esse tipo de tarefa difere da classificação por causa do formato da saída (GOODFELLOW; BENGIO; COURVILLE, 2016). A regressão é vista como um caso especializado de classificação para o qual as variáveis de saída são contínuas em vez de categóricas. (NICOLAS, 2017)

Desta forma, podemos afirmar que a regressão é uma técnica de classificação particularmente adequada para um modelo contínuo. As regressões lineares, polinomiais e logísticas estão entre as técnicas mais usadas para ajustar um modelo ou função paramétrica, em um conjunto de dados (NICOLAS, 2017). Em outras palavras, a regressão é um algoritmo de aprendizado de máquina que pode ser treinado para prever saídas numeradas reais, o que pode ser aplicado na área de saúde, como por exemplo, para prever os futuros estados de saúde dos pa-

cientes e identificar pacientes de alto risco para doenças específicas, onde o risco é uma função da probabilidade de precisar atendimento de emergência (REDDY; AGGARWAL, 2015).

2.2.3 LSTM

As redes de memória de longo prazo (LSTM - *Long short-term memory*) são um tipo especial de rede neural recorrente (RNN), capaz de aprender dependências de longo prazo. Foram introduzidos por (HOCHREITER; SCHMIDHUBER, 1997) e, ao invés de uma unidade que simplesmente aplica uma não linearidade por elementos à transformação afim de entradas e unidades recorrentes, as redes recorrentes LSTM têm “células LSTM” que possuem uma recorrência interna (um auto-loop), além da recorrência externa da RNN. Cada célula possui as mesmas entradas e saídas que uma rede recorrente comum, mas possui mais parâmetros e um sistema de unidades de controle que controla o fluxo de informações (GOODFELLOW; BENGIO; COURVILLE, 2016).

A ideia de introduzir *auto-loops* para produzir caminhos nos quais o gradiente pode fluir por longas durações é uma das principais contribuições do modelo inicial de memória de longo prazo (LSTM) (HOCHREITER; SCHMIDHUBER, 1997). Ao aumentar o peso desse *auto-loop* que é controlado por uma unidade oculta (GERS; SCHMIDHUBER; CUMMINS, 1999), a escala de tempo da integração pode ser alterada dinamicamente. Em outras palavras, podemos dizer que mesmo para uma LSTM de parâmetros fixos, a escala de tempo da integração pode mudar com base na sequência de entrada, porque as constantes de tempo são produzidas pelo próprio modelo (GOODFELLOW; BENGIO; COURVILLE, 2016).

Uma unidade LSTM comum é composta por uma célula, uma *input gate* (porta de entrada), uma *output gate* (porta de saída) e uma *forget gate* (porta de esquecer). A célula lembra valores em intervalos de tempo arbitrários, e as três portas regulam o fluxo de informações para dentro e para fora da célula (JUNWEI et al., 2019). A Figura 10 mostra a estrutura de uma célula LSTM tradicional e ilustra as operações das portas *input*, *forget* e *output* na célula básica do LSTM, e cada porta possui uma função de ativação sigmóide e uma operação de multiplicação por pontos. A célula básica do LSTM é definida com as seguintes equações:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2.1)$$

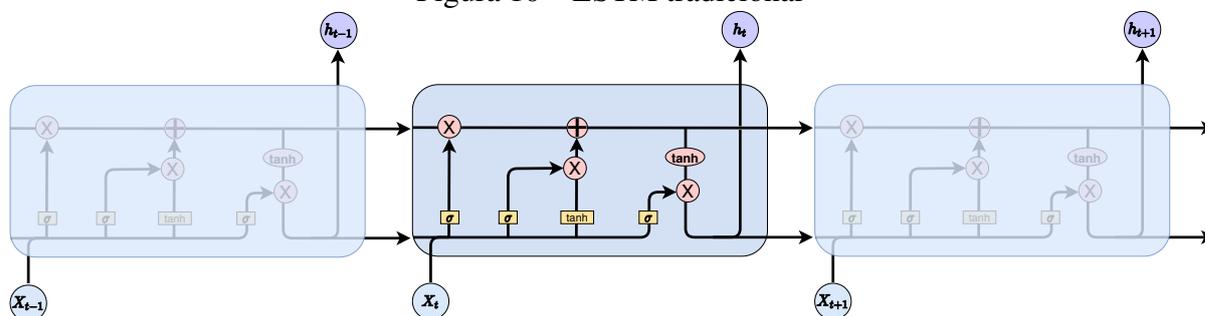
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (2.2)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (2.3)$$

onde f_t denota a saída do *forget gate* para a rede na etapa t , onde σ é a função sigmóide logística. i_t e o_t denotam a saída do *input gate* e do *output gate*, respectivamente. x_t e h_{t-1} são a entrada e o estado oculto anterior, respectivamente. W_f, W_i, W_o, b_f, b_i e b_o são matrizes de

peso que são aprendidas.

Figura 10 – LSTM tradicional



Fonte: Adaptado de (JUNWEI et al., 2019)

Foi demonstrado que as redes LSTM aprendem dependências de longo prazo de forma mais eficiente do que as arquiteturas recorrentes simples (GOODFELLOW; BENGIO; COURVILLE, 2016). Desta forma, diversos trabalhos na área de saúde tem demonstrado o uso de redes LSTM para processar e prever eventos importantes com intervalos e atrasos relativamente longos em séries temporais (JUNWEI et al., 2019; MASSARO et al., 2019; FARUQUI et al., 2019; YIN et al., 2019; NGUYEN; ROKICKI, 2018).

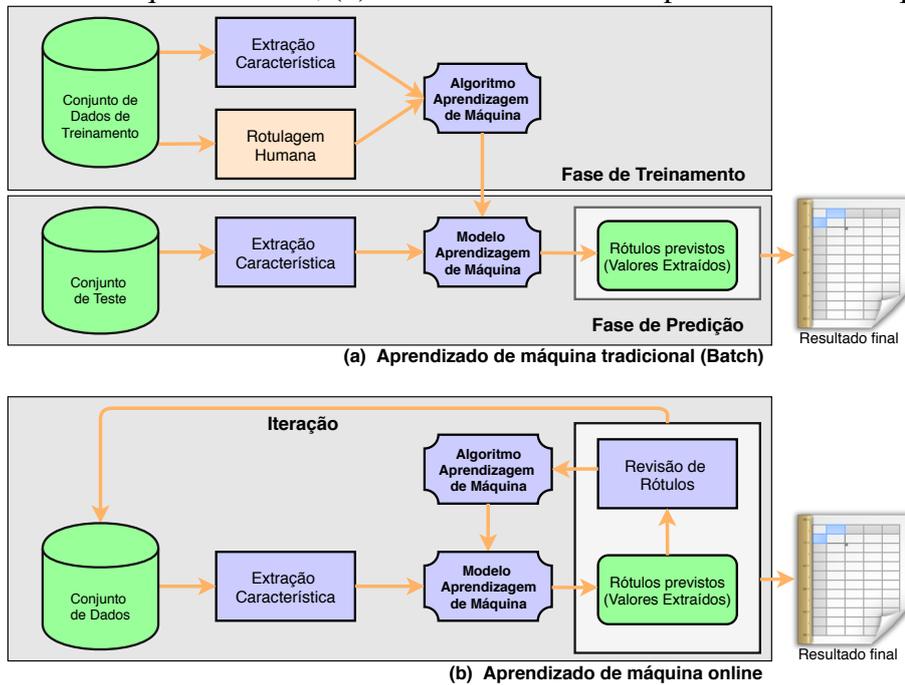
2.2.4 Aprendizagem de máquina online

Tradicionalmente o aprendizado de máquina é executado em modo *batch* (lote). Um exemplo disso seria uma tarefa de aprendizado supervisionado, em que uma coleta de dados de treinamento é fornecida antecipadamente para treinar um modelo seguindo um algoritmo de aprendizado. Essa abordagem exige que todo o conjunto de dados de treinamento seja disponibilizado antes da tarefa de aprendizado, e o processo de treinamento geralmente é realizado em um ambiente offline (HOI et al., 2018). Essa abordagem possui alguns problemas críticos: a) baixa eficiência nos custos de tempo e espaço; e b) baixa escalabilidade, porque o modelo geralmente precisa ser re-treinado do zero para novos dados de treinamento (LOSING et al., 2019; HOI et al., 2018).

Conforme pode ser visto na Figura 11(a), os algoritmos tradicionais de aprendizado de máquina adotam uma abordagem de dois estágios: treinamento em lote (batch) com base em um conjunto de dados de treinamento e previsão para conjunto de dados futuros, com base no modelo gerado a partir do estágio de treinamento. Por sua vez, os algoritmos de aprendizado de máquina online adotam uma abordagem iterativa, conforme pode ser visto na Figura 11(b), onde ele aprende com um dado por vez, e prevê valores a serem extraídos para o próximo dado. O aprendizado irá ocorrer a partir das revisões feitas pelo usuário, e o modelo atualizado é aplicado para previsão dos dados subsequentes (ZHENG et al., 2017).

A configuração de aprendizado online é mais desafiadora, pois os dados são acessados um por um em uma ordem predefinida e o algoritmo fornece um modelo após cada ponto de da-

Figura 11 – Aprendizado de máquina online versus aprendizado em lote. (a) fluxo de trabalho de aprendizado de máquina em lote; (b) Fluxo de trabalho de aprendizado de máquina online.



Fonte: Adaptado de (ZHENG et al., 2017)

dos. Inicialmente, os algoritmos online tendem a oferecer um desempenho mais baixo em comparação com seus equivalentes offline. entretanto, possuem os benefícios de uma baixa complexidade de tempo e espaço, são capazes de processar conjuntos de dados de tamanhos arbitrários e permitem um ajuste específico a um domínio de problema (LOSING et al., 2019). Formalmente, uma sequência potencialmente infinita $S_t = (s_1, s_2, \dots, s_t)$ de tuplas $s_i = (x_i, y_i)$ chega uma após a outra. Ao contrário da configuração offline, uma função de modelo é gerada após cada tupla. Como t representa o registro de data e hora atual, a classificação $\hat{y}_t = h_{t-1}(x_t)$ é feita de acordo com o modelo aprendido anteriormente h_{t-1} . Após o rótulo verdadeiro y_t ser revelado, o algoritmo de aprendizado aplicado gera um novo modelo $h_t = \text{train}(h_{t-1}, s_t)$ com base na tupla atual s_t e no modelo anterior h_{t-1} . Normalmente, o erro do treino de teste intercalado é usado para avaliação de desempenho e é definido como (LOSING et al., 2019):

$$\hat{E}(S_t) = \frac{1}{t} \sum_{i=1}^t \ell(h_{i-1}(x_i), y_i) \quad (2.4)$$

Na área de saúde, sistemas de sensores podem gerar uma quantidade de dados muito grande. Como as técnicas de aprendizado de máquina online treinam seu classificador quando uma nova instância de dados está disponível em tempo de execução (REDDY; AGGARWAL, 2015), podemos pensar em algoritmos de aprendizado online implementados em um sistema de processamento de fluxo de dados de leitura de sensores, aplicados ao monitoramento da saúde pacientes com doenças crônicas, por exemplo.

2.3 Considerações parciais

Um dos maiores desafios enfrentados pela prática médica é atender as necessidades de pacientes com doenças ou comprometimentos crônicos, como o DM2. Em geral, pacientes com doenças crônicas não estão recebendo terapias eficazes e tem pouco controle da doença (WAGNER, 1998). Nas doenças crônicas, o sucesso do tratamento depende de alguns fatores importantes, como o autocuidado apoiado, suporte à decisões clínicas e sistemas de informações clínicas aprimorados (Brasil, Ministério da Saúde, 2013).

A inteligência artificial (IA) tem sido aplicada em vários campos médicos para muitos propósitos (HAMET; TREMBLAY, 2017). Métodos de IA em combinação com tecnologias mais recentes, incluindo a computação móvel e tecnologias de sensores, tem a capacidade de permitir a criação e prestação de serviços de gerenciamento de saúde melhores (CONTRERAS; VEHI, 2018). Um dos desafios da medicina associada a computação móvel e sensores, é a geração de grandes quantidades de dados, para isso, técnicas de aprendizado de máquina tem sido utilizados para a análise e aplicação de conhecimento obtido desta grande quantidade de dados (HAMET; TREMBLAY, 2017).

Dentro desta perspectiva, diversos trabalhos que aplicam métodos de inteligência artificial em diabetes mellitus tipo 2 tem sido realizados. Estes trabalhos são aplicados principalmente para diagnóstico, tomada de decisão terapêutica e previsão de resultados, com destaque para a triagem e diagnóstico da DM2 utilizando técnicas de aprendizado de máquina (ABHARI et al., 2019). Na prática médica de rotina, a IA na medicina tem sido associada ao desenvolvimento de programas para ajudar os médicos a formular diagnósticos e previsão de status crítico, como uma emergência ou piora da condição do paciente, em suas atividades diárias (RIGLA et al., 2018).

O gerenciamento ideal do DM2 exige a reformulação do sistema atual de prestação de assistência médica, mudando o foco de atendimento reativo para atendimento proativo. A medicina preditiva e preventiva para DM2 deve contar com a capacidade de capitalizar informações de uma ampla gama de dados (estilo de vida, social, clínico, tratamento e molecular), a fim de detectar precocemente alterações fisiopatológicas e melhorar a intervenção e o tratamento individualizado (ZARKOGIANNI et al., 2015).

3 TRABALHOS RELACIONADOS

Com o avanço e a aplicação da tecnologia computacional, a computação onipresente agora é usada em todos os lugares (COSTA; YAMIN; GEYER, 2008). Com a miniaturização de dispositivos incorporados, bem como com a natureza cada vez mais inteligente do software, a interação homem-computador está aumentando continuamente. Além dessa interação, o hardware e o software devem cooperar, melhorando a interação homem-máquina. Uma das características críticas da computação ubíqua é a ciência de contexto (DEY; ABOWD; SALBER, 2001; TAVARES et al., 2016). A ideia da ciência de contexto é fornecer os serviços apropriados aos usuários, como residências inteligentes, escritórios inteligentes e serviços relacionados à saúde (MSHALI; LEMLOUMA; MAGONI, 2018; MONTORI et al., 2018). Além disso, atualmente a Internet das Coisas (IoT) vem emergindo como um novo paradigma em tecnologia da informação. A ideia central é construir uma infraestrutura de rede dinâmica, conectando uma variedade de “coisas” físicas e virtuais ao uso de dispositivos e sensores móveis. O uso prático desses sensores em uma plataforma conectada à Internet gera muitas possibilidades de pesquisa, como arquitetura do sistema e processamento de dados de aplicativos (ČOLAKOVIĆ; HADŽIALIĆ, 2018; FIROUZI et al., 2018). Colaborando com isso, a tecnologia da Internet das Coisas conduziu a assistência médica através de uma rápida proliferação de dispositivos portáteis e smartphones, de um sistema convencional baseado em hub a sistemas de assistência médica mais personalizados. A tecnologia da IoT em sistemas personalizados ainda é um desafio significativo nesse campo, sofrendo de problemas como a falta de sensores médicos precisos e econômicos, arquiteturas não padronizadas do sistema da IoT, heterogeneidade de dispositivos vestíveis conectados, multidimensionalidade dos dados e a alta demanda por interoperabilidade (BALOCH; SHAIKH; UNAR, 2018; QI et al., 2017; PASQUIER et al., 2018).

Tradicionalmente, a motivação para usar as tecnologias da informação e comunicação (TIC) em um sistema de saúde é fornecer serviços médicos aos pacientes, chamados e-health, como sistemas de registro eletrônico, sistemas de telemedicina e dispositivos para diagnóstico (QI et al., 2017). Porém, com o aumento da longevidade e o aumento do número de pessoas com doenças crônicas, há problemas com a viabilidade econômica dos sistemas de saúde tradicionais, gerando a necessidade de desenvolver sistemas de saúde onipresentes para fornecer serviços de saúde de qualidade centrados no paciente (BOBEK; NALEPA, 2017). Isso inclui, por exemplo, o desenvolvimento de sistemas de saúde eletrônica aplicados a casas inteligentes (MSHALI; LEMLOUMA; MAGONI, 2018), sistemas de monitoramento generalizado da saúde (AHMADI et al., 2018; ELHELW et al., 2009; REY et al., 2018) e sistemas para serviços de bem-estar de idosos em casas inteligentes (JUNG, 2017). Para a funcionalidade completa desses sistemas, devemos considerar o contexto, que é definido como qualquer informação que possa ser usada para caracterizar a situação de uma pessoa, local ou objeto considerado relevante para a interação entre um usuário e um aplicativo, onde a relevância depende da tarefa do usuário. Como consequência, algumas características desejáveis para sistemas sensíveis ao

contexto em saúde são a apresentação de informações e serviços a um paciente, a execução automática de uma função ou a marcação de contexto para fins de recuperação posterior (DEY; ABOWD; SALBER, 2001). Assim, no campo da saúde, sensores vestíveis, como acelerômetros ou ECG, são acoplados ao paciente para medir dados fisiológicos e ambientais. Esses dados são então analisados para fornecer feedback para diversos propósitos, como avaliar a eficácia de um novo tratamento, estudar melhor uma doença, superar o comportamento do paciente ou ajustar o medicamento. Os aplicativos de assistência médica sensíveis ao contexto enfrentam alguns desafios comuns, independentemente dos objetivos finais (RAULT et al., 2017; GARRIDO; PENICHER; LOZANO, 2018).

Como exemplos de desafios relacionados ao uso da ciência de contexto em saúde, podemos citar a integração de dados heterogêneos de sensores, com diferentes tecnologias, incluídas nessas redes e protocolos (ESPOSITO et al., 2018; De La Iglesia et al., 2018). A eficiência energética dos sensores, bem como sua capacidade de resposta e robustez, são alguns dos desafios que surgem (MICHALAKIS; CARIDAKIS, 2017; RAHMANI et al., 2018; DEEN, 2015). Outro desafio é a alternância de contexto e entrega contínua de dados, sem intervalos (BALLOCH; SHAIKH; UNAR, 2018; BOBEK; NALEPA, 2017). Há também questões relacionadas à segurança e privacidade das informações de saúde, importantes na construção do contexto em que o paciente está sob cuidados (WETZELS et al., 2018; AMIN et al., 2018; KUIJS; ROSENCRANTZ; REICH, 2015). Finalmente, o crescente uso da inteligência artificial e aprendizado de máquina na área da saúde tem aparecido como uma tendência significativa (HOLZINGER et al., 2017), embora se depare com o dilema entre vigilância e controle, por um lado, e apoio e benefícios pessoais, por outro (RÖCKER; ZIEFLE; HOLZINGER, 2014).

Esse capítulo apresenta uma revisão de literatura que contempla trabalhos no que tange ao conceito de ciência de contexto especificamente relacionado a saúde e discute alguns problemas e desafios que cercam a área. A motivação da revisão ter se dado no contexto da saúde em uma perspectiva ampla vem ao encontro do fato de a diabetes mellitus tipo 2 ser uma doença que está associada a uma série de outras comorbidades (DEFRONZO et al., 2015), ampliando-se assim a abrangência de estudo no que se refere a ciência de contexto. Como resultado dessa revisão de literatura, encontram-se diversos desafios a respeito de: a) aquisição de contexto do paciente; b) modelagem e integridade dos dados e c) métodos de fusão de dados e raciocínio. Dessa forma, os desafios destacados puderam ser mapeados à realidade do paciente crônico de diabetes mellitus tipo 2, fornecendo assim informações relevantes para orientar a elaboração desta proposta de tese.

O capítulo está organizado da seguinte forma, a seção 3.1 descreve a metodologia adotada para elaborar a revisão da literatura. A seção 3.2 apresenta os resultados da revisão da literatura com os artigos selecionados. A seção 3.4 discute os artigos que destacam os principais desafios e oportunidades de pesquisa. Por fim, a seção 3.5 estabelece uma conexão entre a ciência de contexto na saúde e os desafios das doenças crônicas, particularmente da diabetes mellitus tipo 2.

3.1 Método

Esta seção foca na descrição do protocolo do estudo, que apresenta os procedimentos adotados e orienta as decisões subsequentes. Ao contrário de um processo de revisão convencional e não estruturado, uma revisão sistemática segue uma sequência rigorosa de etapas metodológicas (BRERETON et al., 2007; ZHANG; Ali Babar, 2013). Com base em um protocolo de revisão bem definido e avaliado, este estudo apresenta uma revisão sistemática da literatura, projetada para fornecer uma visão mais ampla da conscientização do contexto na pesquisa em saúde.

3.1.1 Projeto de estudo

Optamos pela abordagem sistemática de revisão de literatura, pois um grande número de artigos pode ser combinado e, ao mesmo tempo, pode fornecer um resumo, permitindo identificar lacunas para pesquisas futuras. A revisão sistemática da literatura é cada vez mais utilizada no campo da saúde, tema central deste estudo, fator que contribuiu para sua adoção. As etapas abaixo definem o escopo desta revisão sistemática de literatura (RSL) (ROEHRS et al., 2017):

1. Questões de pesquisa: apresenta as questões de pesquisa utilizadas;
2. Estratégia de pesquisa: apresenta a estratégia e as bibliotecas exploradas para coletar os dados;
3. Seleção de artigo: explica os critérios para seleção dos estudos;
4. Distribuição dos artigos: apresenta como os estudos são distribuídos cronologicamente;
5. Avaliação da qualidade: avalia a qualidade dos estudos selecionados;
6. Extração de dados: compara os estudos selecionados com as questões de pesquisa.

3.1.2 Questões de pesquisa

As perguntas de pesquisa são baseadas na motivação para conduzir a revisão sistemática, ou seja, as respostas a essas perguntas devem fornecer uma consolidação baseada em evidências para definir, aplicar e adquirir conhecimento sobre o uso da conscientização do contexto na área da saúde.

De acordo com a tabela 2, a pergunta RQ1 diz respeito às definições, classificações e caracterizações existentes na literatura (se houver). A RQ2 refere-se aos desafios que surgem no uso de informações sensíveis ao contexto no campo da saúde. A pergunta RQ3 refere-se a cenários nos quais dados sensíveis ao contexto são aplicados ou seu uso tem relevância para o campo da saúde. A pergunta RQ4 procura identificar e comparar as evidências relacionadas aos métodos de modelagem de saúde com reconhecimento de contexto. Finalmente, a pergunta RQ5

pergunta como o raciocínio contextual é apresentado, incluindo também métodos de fusão de dados.

Tabela 2 – Questões de pesquisa

Identificador	Questão
RQ1	Como é definido e classificado o uso da ciência de contexto no campo da saúde na pesquisa existente?
RQ2	Quais são os desafios relacionados ao uso de informações sensíveis ao contexto na área da saúde?
RQ3	Quais são os cenários de casos de uso relacionados à ciência de contexto na área da saúde?
RQ4	Quais são as técnicas de modelagem de contexto empregadas na área da saúde?
RQ5	Quais são os métodos de raciocínio e fusão de dados utilizados em informações sensíveis ao contexto na área da saúde?

Fonte: Elaborado pelo Autor

3.1.3 Estratégia de pesquisa

Após definir as questões de pesquisa, o próximo passo foi selecionar um conjunto de estudos relacionados às questões de pesquisa. Esse processo envolve a definição de palavras-chave de pesquisa e a definição do escopo da pesquisa (PETTICREW; ROBERTS, 2006). As palavras-chave de pesquisa foram definidas após a leitura de artigos na área de interesse e a separação de termos, sinônimos e acrônimos que melhor definiram o objeto de pesquisa (levando em consideração sua relação com questões de pesquisa definidas anteriormente). Seguindo o que foi definido por Kitchenham e Charters (KITCHENHAM, 2004), esses termos foram combinados usando operadores booleanos.

Definimos o escopo do trabalho como o uso de dados sensíveis ao contexto no campo da saúde. Dessa forma, a cadeia final de pesquisa, formada a partir dos critérios descritos acima, foi definida a seguir.

String de pesquisa

(Context aware OR Context awareness OR Situational Aware OR Situational Awareness)
AND (Health OR HealthCare) AND (Environment data)

Na execução da busca de artigos, os estudos foram obtidos a partir de bases de dados eletrônicas, selecionadas através de buscas utilizando as palavras-chave da string construída. Os bancos de dados eletrônicos incluídos na pesquisa foram: ACM Digital Library¹; Google scholar²;

¹<https://dl.acm.org/>

²<https://scholar.google.com.br/>

IEEE Xplore Library³; IET Digital Library⁴; JMIR Publications⁵; PubMed⁶; ScienceDirect⁷; Springer Link⁸; Web of Science⁹; and Wiley Online Library¹⁰.

Esses bancos de dados foram escolhidos porque consistem em uma amostra significativa das bases de dados e fornecem textos completos e *proceedings* das conferências mais importantes na área de saúde, que envolvem e-health, wearables e suas relações.

3.1.4 Seleção de Artigos

Após coletar o conjunto de artigos relacionados nas bases de dados, seguiu-se a remoção de estudos que não eram relevantes, de forma a manter apenas os mais representativos. Para isso, após a pesquisa inicial, as impurezas foram removidas, o que implicou a remoção de artigos duplicados (porque alguns estudos estavam disponíveis em mais de um banco de dados) e artigos não relacionados à string de pesquisa, mas que retornaram devido a características dos diferentes bancos de dados eletrônicos.

Após o primeiro filtro, o segundo e o terceiro filtros consistiram em analisar o título e o resumo, respectivamente, e excluir os que não mencionavam saúde e contexto (de acordo com a string de pesquisa). Na quarta etapa, os demais estudos foram agrupados e os capítulos de livros, teses e artigos curtos foram removidos. Com relação à retirada dos artigos curtos, foi utilizado o critério de possuir no mínimo cinco páginas, e alguns artigos com menos de cinco páginas foram mantidos por apresentarem algumas características desejáveis, como trabalhos relacionados e resultados finais. Por fim, alguns estudos não relacionados a esta pesquisa permaneceram. Analisamos o texto completo para remover aqueles que não eram relevantes.

3.1.5 Avaliação da qualidade

Uma das preocupações da revisão é avaliar a qualidade do corpus selecionado. De acordo com (KITCHENHAM, 2004), o critério de qualidade verifica se o artigo é um estudo significativo. Para tanto, foram elaboradas perguntas, as quais foram utilizadas como critério para avaliar os artigos encontrados. De acordo com a Tabela 3, verificou-se a presença dos seguintes itens em cada artigo: proposta de pesquisa, revisão da literatura, trabalhos relacionados, métodos, resultados, conclusões e trabalhos futuros.

³<https://ieeexplore.ieee.org/Xplore/home.jsp>

⁴<http://digital-library.theiet.org/>

⁵<http://jmirpublications.com/>

⁶<https://www.ncbi.nlm.nih.gov/pubmed/>

⁷<https://www.sciencedirect.com/>

⁸<https://link.springer.com/>

⁹<http://isiknowledge.com>

¹⁰<https://onlinelibrary.wiley.com/>

Tabela 3 – Critérios de avaliação da qualidade

Critério	Descrição
Critério 1	O artigo tem uma proposta de pesquisa.
Critério 2	O artigo apresenta uma revisão de literatura.
Critério 3	O artigo discute trabalhos relacionados.
Critério 4	O artigo possui uma metodologia.
Critério 5	O artigo apresenta resultados.
Critério 6	O artigo tem uma conclusão.
Critério 7	O artigo sugere pesquisas futuras.

Fonte: Elaborado pelo autor

3.1.6 Extração de dados

A extração e síntese dos dados foram realizadas através da leitura de cada um dos artigos incluídos na revisão e da extração dos dados relevantes. Isso foi gerenciado através da ferramenta de gerenciamento bibliográfico Mendeley e software de planilha de dados. Para manter as informações consistentes, a extração de dados para os estudos incluídos na revisão foi baseada nas perguntas da pesquisa e onde elas foram respondidas no corpo do artigo, como pode ser visto na Tabela 4. Para sintetizar os dados, inspecionamos os dados extraídos em busca de similaridades. Os resultados da síntese serão descritos nas seções subsequentes.

Tabela 4 – Revisão dos artigos relacionados às perguntas da pesquisa

Section	Description	Research Questions
Conteúdo Aberto		
	Título	RQ4, RQ5
	Resumo	RQ1, RQ3, RQ5
	Palavras-chave	RQ1, RQ3, RQ4, RQ5
Conteúdo do Artigo		
	Introduction	Todas as perguntas
	Metodologia	Todas as perguntas
	Resultados	Todas as perguntas
	Discussão	Todas as perguntas
	Conclusão	Todas as perguntas

Fonte: Elaborado pelo autor

3.2 Resultados

Nesta seção, apresentamos os resultados dos estudos avaliados, relacionados ao tema da pesquisa. Nas subseções a seguir, tentamos responder a cada pergunta de pesquisa proposta sintetizando as informações. Como resultado, além de responder às perguntas da pesquisa,

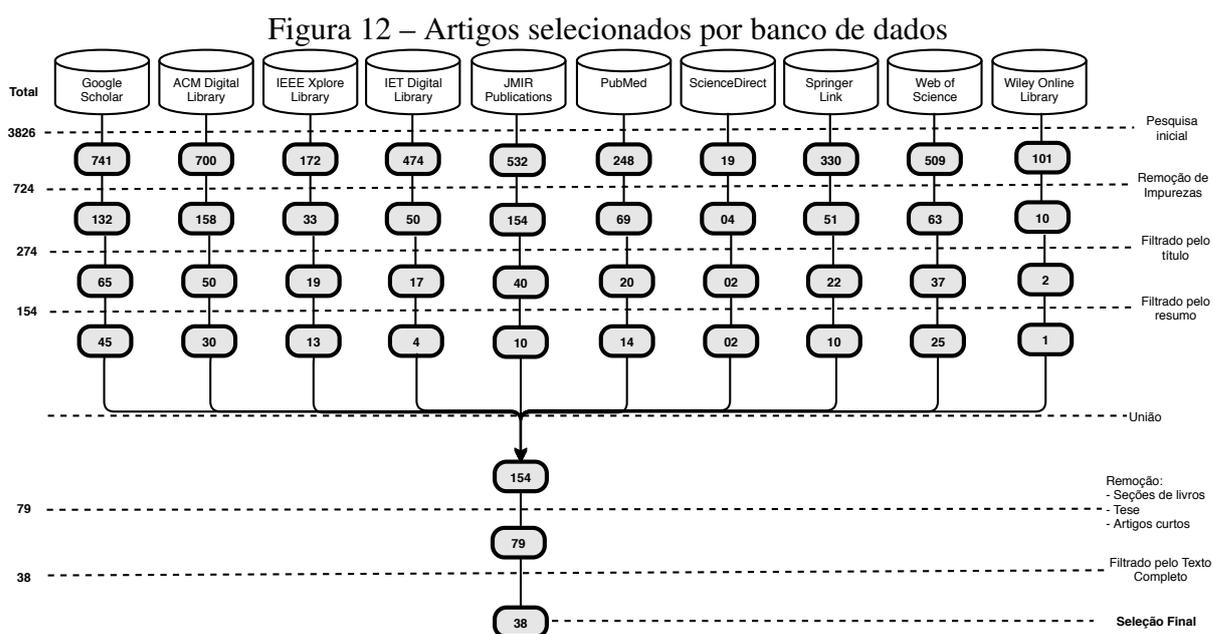
também são sugeridas contribuições no campo do uso de dados sensíveis ao contexto no campo da saúde, com uma taxonomia proposta e uma visão atualizada dos principais desafios.

3.2.1 Estratégia para conduzir a pesquisa

Para esta pesquisa, escolhemos dez bancos de dados eletrônicos (listados na seção 3.1.3) para selecionar os artigos. Os critérios para a escolha desses bancos de dados foram a cobertura dos campos da ciência da computação e saúde. Para a seleção de artigos em cada banco de dados, foram utilizados alguns procedimentos para limitar nossa pesquisa, selecionando artigos em inglês, no período de 2008 a 2018, e excluindo resultados de patentes e citações.

3.2.2 Prosseguindo com a seleção de artigos

Nosso processo de busca é apresentado em detalhes na Figura 12, mostrando os processos de exclusão e filtragem.



Fonte: Elaborado pelo autor

Inicialmente, nossa string de pesquisa encontrou 3.826 artigos nos diferentes bancos de dados. Após o resultado inicial, utilizamos o critério de remoção de impurezas, que consiste na remoção de artigos publicados em periódicos e conferências não relacionados ao tema desta pesquisa e na remoção de artigos duplicados, restando 724 estudos. Após o primeiro filtro, usamos o título e o resumo do artigo como critério de exclusão. Foram excluídos do corpus os estudos em que os títulos ou o resumo não apresentavam palavras contidas na *string* de busca. A inclusão de artigos que não apresentavam estas palavras pode ter ocorrido devido às diversas alterações efetuadas na *string*, para adaptá-la à pesquisa efetuada nas bases em que

foram efetuadas as buscas. Restando 154 artigos que foram agrupados para serem aplicados novos critérios de exclusão.

Após o agrupamento dos artigos, os critérios de exclusão aplicados foram a remoção de capítulos de livros, teses e artigos curtos, restando 79 artigos. O critério para artigos curtos foi utilizado para eliminar artigos com cinco ou menos páginas, exceto artigos que apresentavam propostas de arquiteturas e apresentavam resultados finais.

Nos 79 artigos restantes foi realizada uma análise, e os artigos que não contribuíram significativamente para os critérios utilizados nesta pesquisa foram eliminados. Além disso, alguns desses estudos pertenciam ao mesmo grupo de pesquisadores, apresentando os mesmos métodos ou técnicas, dentre os quais foram selecionados apenas os artigos mais representativos. Assim, 38 artigos foram selecionados como base para este estudo. Uma visão geral dos estudos selecionados é apresentada na Tabela 5, mostrando o identificador, referência, ano de publicação, editor e tipo.

Tabela 5: Corpus final de artigos publicados

Artigo	Ano	Editora	Tipo
(BALOCH; SHAIKH; UNAR, 2018)	2018	Springer	Journal
(MSHALI; LEMLOUMA; MAGONI, 2018)	2018	Elsevier	Journal
(MSHALI et al., 2018)	2018	Elsevier	Journal
(NEWCOMBE et al., 2018)	2018	IEEE	Conference
(ALIREZAIE et al., 2017)	2017	MDPI	Journal
(AZIMI et al., 2017)	2017	Springer	Journal
(GOMES et al., 2017)	2017	Wiley	Journal
(JUNG, 2017)	2017	MDPI	Journal
(MICHALAKIS; CARIDAKIS, 2017)	2017	Springer	Journal
(MATTHEW-MAICH et al., 2016)	2016	JMIR	Journal
(SANTOS et al., 2016)	2016	Springer	Journal
(CHIANG; LIANG, 2015)	2015	Springer	Journal
(DEEN, 2015)	2015	Springer	Journal
(DOBRESCU, 2015)	2015	Inase	Journal
(ELMALAKI; WANNER; SRIVASTAVA, 2015)	2015	ACM	Journal
(FORKAN et al., 2015)	2015	IEEE	Journal
(HAMEURLAINE et al., 2015)	2015	Springer	Journal
(KUIJS; ROSENCRA NTZ; REICH, 2015)	2015	Iaria	Conference
(ALSHURAF A et al., 2014)	2014	IEEE	Journal
(BEATTIE et al., 2014)	2014	ACM	Journal
(CHERKA OUI; AGOULMINE, 2015)	2014	IEEE	Journal

Continua na próxima página

Tabela 5 – continuação da página anterior

Artigo	Ano	Editora	Tipo
(MCHEICK, 2014)	2014	Elsevier	Journal
(SANNINO; De Pietro, 2014)	2014	Scopus	Journal
(THOMAS et al., 2014)	2014	Inderscience	Journal
(ZERKOUK et al., 2014)	2014	Springer	Journal
(MACHADO et al., 2013)	2013	Springer	Conference
(MNATSAKANYAN et al., 2012)	2012	Elsevier	Journal
(NAVA-MUÑOZ; MORÁN, 2012)	2012	MDPI	Journal
(O'DONOGHUE; HERBERT, 2012)	2012	ACM	Journal
(SANCHEZ-PI; CARB, 2012)	2012	Springer	Journal
(SONG et al., 2012)	2012	Taylor&Francis	Conference
(BAE et al., 2012)	2012	IEEE	Conference
(MITCHELL et al., 2011)	2011	IEEE	Journal
(TREBATOSKI et al., 2010)	2010	Journals@UIC	Journal
(AL-NEYADI; ABAWAJY, 2009)	2009	Springer	Conference
(ELHELW et al., 2009)	2009	IEEE	Conference
(MILEO et al., 2010)	2009	Oxford	Journal
(KOMNAKOS et al., 2008)	2008	ACM	Journal

Fonte: Elaborado pelo autor

3.2.3 Avaliando a qualidade

Nesta seção, apresentamos uma análise qualitativa dos artigos, com base nos critérios definidos na Seção 3.1.5. Utilizamos uma escala de 0 a 7, como em Roehrs (ROEHRS et al., 2017) para analisar e classificar esses artigos, descritos na Tabela 3, com 31 artigos (81%) cobrindo cinco ou mais dos critérios citados e 7 artigos que cobrem quatro dos critérios de qualidade.

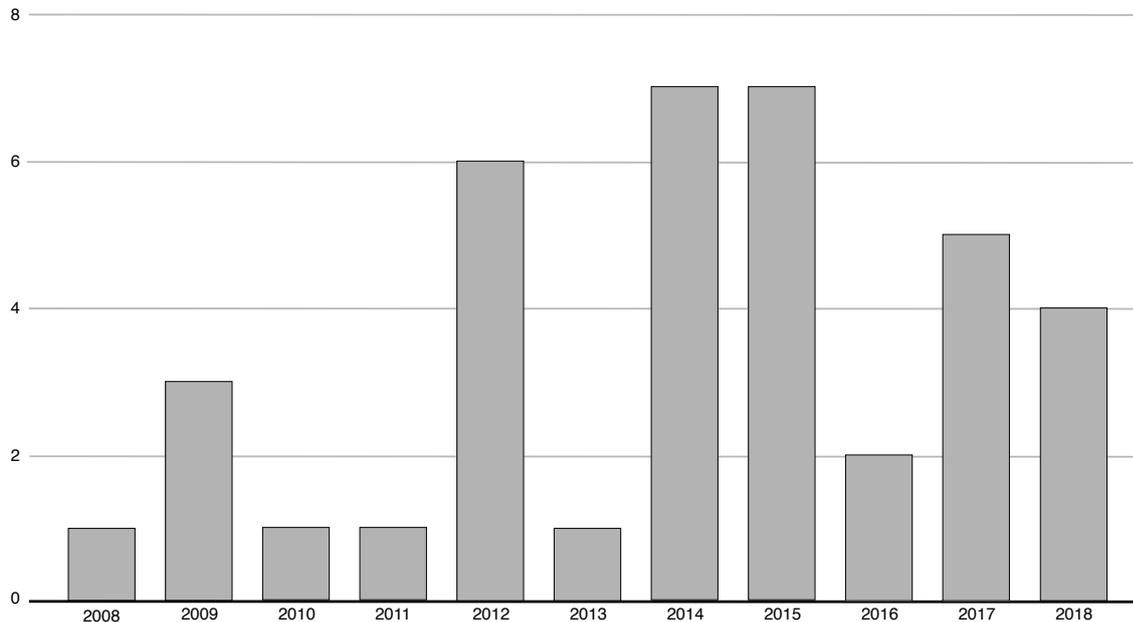
Em relação ao ano de publicação, a distribuição dos artigos é apresentada na Figura 13, dos quais 26 artigos do corpus foram publicados nos últimos cinco anos.

A figura 14 mostra a distribuição do número de citações por artigo. A partir da figura, é possível perceber que os artigos que não possuem citações são em sua maioria a partir do ano de 2018. Por outro lado, os artigos mais citados (mais de 30) foram publicados, em grande parte, nos últimos anos. Os demais artigos possuem, em sua maioria, mais de oito citações.

3.3 Respostas às perguntas e discussão

Esta seção foca nas questões de pesquisa definidas na Seção 3.1.2 para revisar os resultados da revisão de literatura. A literatura pesquisada discute várias definições e classificações, que

Figura 13 – Número de artigos por ano de publicação



Fonte: Elaborado pelo autor

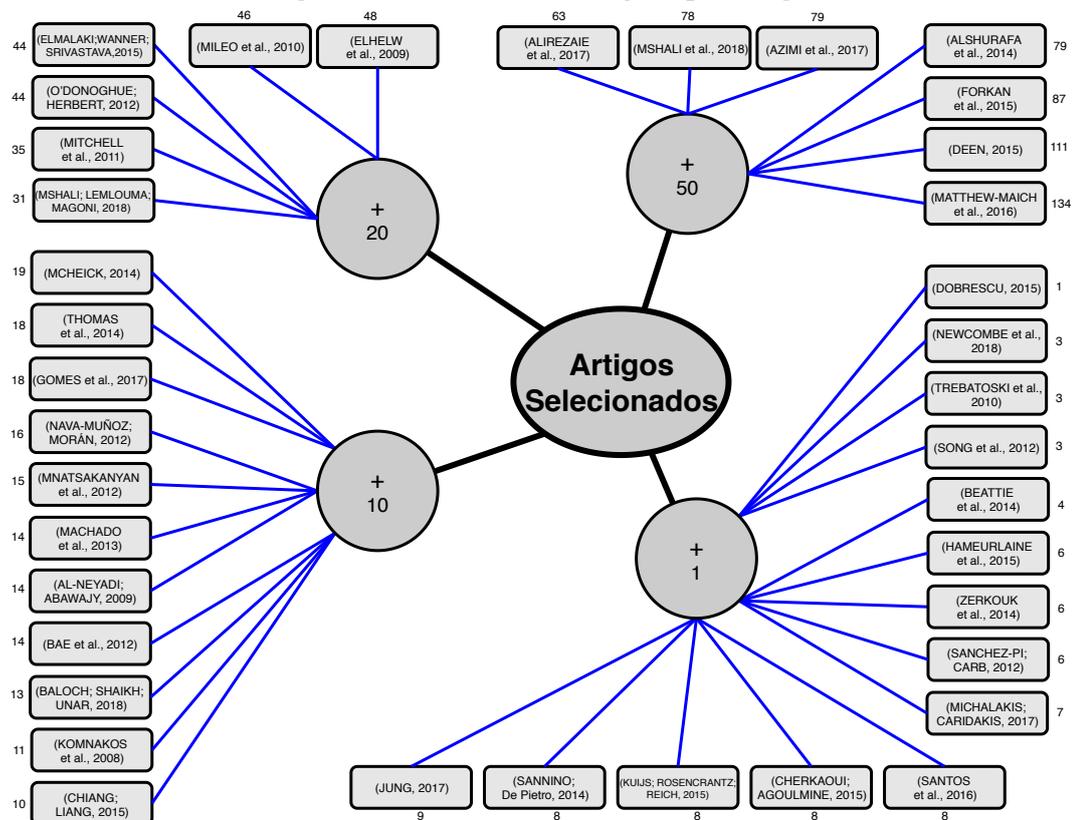
são apresentadas e discutidas abaixo.

3.3.1 Como o uso da ciência de contexto no campo da saúde é definido e classificado nas pesquisas existentes?

A resposta à pergunta RQ1 leva à definição de uma taxonomia baseada na análise dos estudos atuais. Essa taxonomia é ilustrada na Figura 15 e é baseada em observações sobre as tecnologias, métodos de uso e problemas-alvo. Às vezes, as propostas combinam várias técnicas. Como resultado, as classes nesta taxonomia não são mutuamente exclusivas e cada estudo pode ser inserido em uma ou mais delas. O processo de definição da taxonomia começou com uma análise de todos os artigos do corpus para identificar padrões, características e categorias. Além disso, definimos cinco conceitos centrais para analisar (1) sensores, (2) cenários, (3) modelagem de contexto, (4) métodos de fusão e (5) raciocínio de contexto.

O primeiro elemento refere-se aos sensores, que foram classificados em três tipos: físico, virtual e lógico. Os sensores físicos foram separados na taxonomia em sensores incorporados (subdivididos em wearable e interno/smartphones) e sensores externos. Sensores físicos parecem avaliar atividades, como Mshali, Lemlouma e Magoni (MSHALI; LEMLOUMA; MAGONI, 2018), que usam sensores de localização (GPS) e sensores ambientais (umidade, temperatura, câmera) para monitorar as atividades de orientação e memória nos idosos que vivem sozinhos. Outra maneira de usar sensores físicos, desta vez usando um acelerômetro, é incentivar um estilo de vida mais ativo, promovendo exercícios físicos (ALSHURAFI et al., 2014). Diferentemente dos sensores físicos, os sensores virtuais não geram dados por si mesmos, mas

Figura 14 – Número de citações por artigo



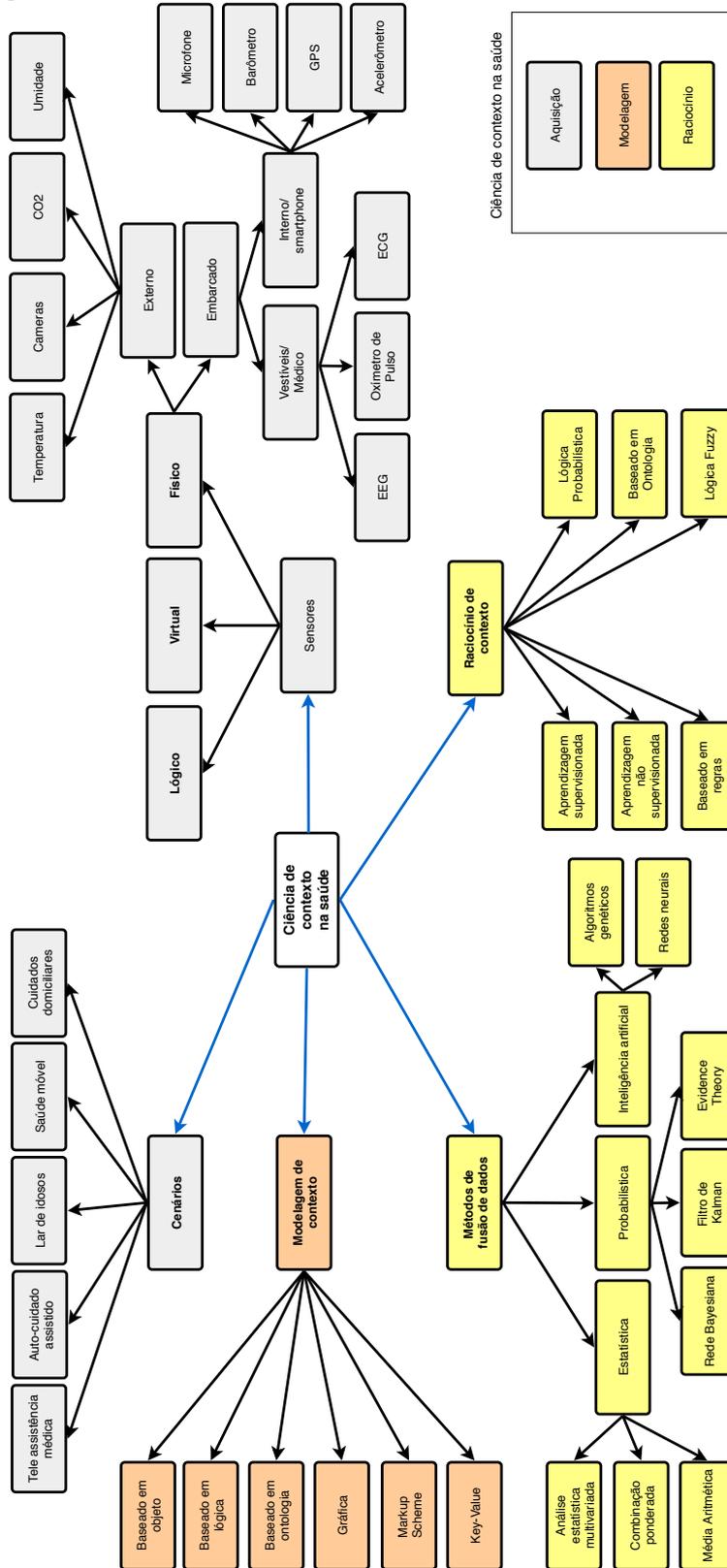
Fonte: Elaborado pelo autor

recuperam dados de diferentes fontes e os publicam como dados do sensor. Os sensores lógicos produzem informações úteis por meio da combinação de sensores físicos e virtuais (por exemplo, um calendário). Em Sannino e De Pietro (SANNINO; De Pietro, 2014), os dados do acelerômetro e da temperatura ambiente (físicos) são combinados com dados de localização e mapas (virtuais) para indicar a localização em situações de risco.

O segundo elemento da taxonomia diz respeito aos cenários em que os dados coletados e o contexto podem auxiliar na assistência à saúde. Alguns pesquisadores propuseram um middleware para a ambientes assistidos (GOMES et al., 2017), aumentando o grau de independência e mobilidade para idosos ou pessoas com doenças crônicas. Chiang e Liang (CHIANG; LIANG, 2015) propõem um sistema inteligente de atendimento domiciliar, que também é aplicado a idosos e pacientes com doenças crônicas. Mshali, Lemlouma e Magoni (MSHALI; LEMLOUMA; MAGONI, 2018) sugerem um sistema de eSaúde adaptável para casas inteligentes, capaz de detectar mudanças de comportamento, concentrando-se nos idosos e nas pessoas que vivem sozinhas. Nava-Muñoz e Morán (NAVA-MUÑOZ; MORÁN, 2012) já relatam a falta de conscientização sobre as situações que envolvem o cuidado ao idoso e a falta de informações sobre a disponibilidade de cuidadores em um lar de idosos, propondo um modelo de notificação com reconhecimento de contexto.

O terceiro elemento da taxonomia refere-se à modelagem de contexto. Os dados devem ser

Figura 15 – Taxonomia do uso de ciência de contexto em saúde



Fonte: Elaborado pelo autor

modelados e representados de acordo com seu significado. Para isso, diferentes técnicas são utilizadas. Bae, Alkobaisi e Narayanappa (BAE et al., 2012) propõem uma estrutura de análise

de dados para apoiar as decisões de saúde ambiental, com modelagem de contexto baseada em ontologia. Alirezaie et al. (ALIREZAIE et al., 2017) apresenta um sistema para casas inteligentes, também com uma técnica de modelagem de contexto baseada em ontologia. Mcheick (MCHEICK, 2014) propõe um formulário híbrido para modelagem de contexto, usando uma abordagem baseada em objetos e marcação utilizando XML.

Os quarto e quinto elementos da taxonomia se referem ao raciocínio de contexto. Este elemento inclui técnicas de fusão de dados e raciocínio. Hameurlaine et al. (HAMEURLAINE et al., 2015) propõe um modelo baseado em regras para raciocinar sobre informações contextuais para fornecer serviços apropriados em sistemas de cuidados de saúde ubíquos. Da mesma forma, Mileo et al. (MILEO et al., 2010) usa um componente de raciocínio que aplica regras lógicas destinadas à interpretação correta de informações contextuais incompletas ou inconsistentes aplicadas à Assistência Médica Domiciliar. Alshurafa et al. (ALSHURAF A et al., 2014) usa uma técnica não supervisionada para agrupamento na proposta de uma estrutura de reconhecimento de atividade física orientada à saúde usando wearables. Elhelw et al. (ELHELW et al., 2009) propõe um método estatístico para mesclar informações de wearables e sensores ambientais aplicados em uma estrutura para monitoramento generalizado de assistência médica e usa uma técnica probabilística para raciocínio de contexto. Mnatsakanyan et al. (MNATSAKANYAN et al., 2012) propõe um modelo de fusão de informações distribuídas com base em redes bayesianas, aplicado à vigilância regional de saúde pública.

3.3.2 Quais são os desafios relacionados ao uso de informações de ciência de contexto na área da saúde?

Para responder à pergunta RQ2, listamos e identificamos desafios comuns no uso de dados de ciência de contexto na área da saúde. Os problemas estão agrupados na Tabela 6. Como pode ser visto, o conteúdo reúne algumas características comuns de desafios relacionados ao uso de tecnologias heterogêneas, desafios associados aos aspectos de comunicação, gerenciamento de energia, escalabilidade e segurança.

Tabela 6: Desafios

Desafio	Artigo
Tecnologias heterogêneas: Integração/Fusão de dados	(ALIREZAIE et al., 2017) (AZIMI et al., 2017) (BAE et al., 2012) (BALOCH; SHAIKH; UNAR, 2018) (CHIANG; LIANG, 2015) (DEEN, 2015) (DOBRESCU, 2015) (ELHELW et al., 2009) (GOMES et al., 2017) (KUIJS; ROSENCRANTZ; REICH, 2015) (MILEO et al., 2010) (MNATSAKANYAN et al., 2012) (MSHALI; LEMLOUMA; MAGONI, 2018) (MSHALI et al., 2018) (O'DONOGHUE; HERBERT, 2012)
Comunicação confiável: Diferentes tecnologias	(ALIREZAIE et al., 2017) (BAE et al., 2012) (CHERKAOUI; AGOULMINE, 2015) (DEEN, 2015) (DOBRESCU, 2015) (ELHELW et al., 2009) (GOMES et al., 2017) (KOMNAKOS et al., 2008) (MATTHEW-MAICH et al., 2016) (MICHALAKIS; CARIDAKIS, 2017) (MILEO et al., 2010) (MSHALI et al., 2018) (NAVA-MUÑOZ; MORÁN, 2012) (O'DONOGHUE; HERBERT, 2012) (SANTOS et al., 2016) (SONG et al., 2012) (ZERKOUK et al., 2014)
Gerenciamento de energia: Consumo de energia	(ALIREZAIE et al., 2017) (BALOCH; SHAIKH; UNAR, 2018) (CHERKAOUI; AGOULMINE, 2015) (DEEN, 2015) (ELHELW et al., 2009) (ELMALAKI; WANNER; SRIVASTAVA, 2015) (FORKAN et al., 2015) (GOMES et al., 2017) (KOMNAKOS et al., 2008) (MITCHELL et al., 2011) (MSHALI et al., 2018) (O'DONOGHUE; HERBERT, 2012) (THOMAS et al., 2014)
Escalabilidade: Funcionalidades de raciocínio e inferência	(ALIREZAIE et al., 2017) (BAE et al., 2012) (BALOCH; SHAIKH; UNAR, 2018) (ELHELW et al., 2009) (GOMES et al., 2017) (KUIJS; ROSENCRANTZ; REICH, 2015) (MSHALI et al., 2018)
Precisão: dados, monitoramento	(BAE et al., 2012) (BALOCH; SHAIKH; UNAR, 2018) (ELHELW et al., 2009) (FORKAN et al., 2015) (KOMNAKOS et al., 2008) (KUIJS; ROSENCRANTZ; REICH, 2015) (MICHALAKIS; CARIDAKIS, 2017) (MILEO et al., 2010) (MITCHELL et al., 2011) (MSHALI; LEMLOUMA; MAGONI, 2018) (MSHALI et al., 2018) (NAVA-MUÑOZ; MORÁN, 2012) (NEWCOMBE et al., 2018) (O'DONOGHUE; HERBERT, 2012) (SANCHEZ-PI; CARB, 2012) (SANNINO; De Pietro, 2014) (THOMAS et al., 2014) (ZERKOUK et al., 2014)

Continua na próxima página

Tabela 6 – continuação da página anterior

Desafio	Artigo
Segurança: Disponibilidade, Privacidade, Confidencialidade	(AL-NEYADI; ABAWAJY, 2009) (ALIREZAIE et al., 2017) (BALOCH; SHAIKH; UNAR, 2018) (FORKAN et al., 2015) (JUNG, 2017) (KUIJS; ROSENCRANTZ; REICH, 2015) (MATTHEW-MAICH et al., 2016) (MICHALAKIS; CARIDAKIS, 2017) (MSHALI et al., 2018) (SANTOS et al., 2016) (ZERKOUK et al., 2014)

Fonte: Elaborado pelo autor

As primeira e segunda linhas da tabela 6 lidam com tecnologias heterogêneas e comunicação confiável. Gomes et al. (GOMES et al., 2017) apontam para a necessidade de uma infraestrutura de software flexível o suficiente para interagir com diferentes tipos de sensores e atuadores e com tecnologias de hardware e comunicação de diferentes formatos. Devemos levar em consideração que os pacotes de dados de diferentes sensores podem ter diferentes formatos e codificações, aumentando a necessidade de módulos de transcodificação que encapsulam a lógica necessária para interpretar os dados trocados (ELHELW et al., 2009).

A terceira linha da tabela 6, Gerenciamento de energia, refere-se ao problema de consumo de energia nos sensores. Assumindo que a maioria dos sensores não possui uma fonte permanente de energia, dependendo das baterias para sua fonte de energia, o consumo se torna crucial para os sensores usados em contextos ambientais e de saúde (GOMES et al., 2017). Além disso, o uso de diferentes tecnologias de transmissão (Zig-Bee, Bluetooth, WiFi) com diferentes necessidades de consumo de transmissão torna o gerenciamento de energia um desafio significativo nesse cenário. Para isso, Alirezaie et al. (ALIREZAIE et al., 2017) faz uso extensivo de protocolos de comunicação de ponta em um sistema sensível ao contexto baseado em ontologia para casas inteligentes, buscando uma comunicação robusta entre dispositivos IoT, mas eficiente em termos de consumo.

As quarta e quinta linhas da Tabela 6, Escalabilidade e Precisão estão relacionadas, no caso de escalabilidade, à capacidade de atender a um número variado e crescente de sensores e, assim, atender ao aumento do processamento relacionado aos dados gerados (GOMES et al., 2017; MSHALI et al., 2018). Por sua vez, precisão e confiabilidade são percebidas como questões críticas em aplicativos sensíveis ao contexto. Às vezes, os sensores vestíveis não fornecem informações suficientes para criar a situação; portanto, são necessários sensores adicionais. Baloch, Shaikh e Unar (BALOCH; SHAIKH; UNAR, 2018) propõem uma abordagem de fusão de dados IoT de saúde em três camadas, que inclui, em um deles, fusão de dados com reconhecimento de contexto, combinando análise de dados de sinais vitais e ambientais para evitar alarmes falsos.

Finalmente, a sexta linha da tabela 6 se refere à segurança. Nos sistemas de monitoramento

de integridade, a segurança é uma questão crítica, pois envolve muitos processos e componentes: sensores e atuadores, coleta de dados e comunicação. Michalakis e Caridakis (MICHALAKIS; CARIDAKIS, 2017) apontam que as pessoas mostram maior sensibilidade ao compartilhar suas informações pessoais de saúde, ou seja, privacidade e segurança são de maior importância para os serviços de saúde em comparação com outros serviços similares. Al-Neyadi e Abawajy (AL-NEYADI; ABAWAJY, 2009) propõem um mecanismo para controlar o acesso aos sistemas de saúde eletrônica com base no contexto. Para garantir que os serviços e as informações sejam acessados apenas por pessoas com privilégios para acessá-los, leva em consideração a pessoa que está tentando acessar os dados, o tipo de dados que está sendo acessado e o contexto da transação na qual a tentativa de acesso esta sendo realizada.

3.3.3 Quais são os cenários relacionados à ciência de contexto na área da saúde?

A resposta à pergunta RQ3 define os cenários nos quais as tecnologias de saúde são aplicadas. Esses cenários são mostrados na Tabela 7 e baseiam-se na análise dos estudos que fazem parte do corpus tratado neste artigo. Geralmente, o mesmo estudo apresenta mais de um cenário de aplicação.

Tabela 7: Cenários

Cenário	Artigo
Cuidados domiciliares	(ALIREZAIIE et al., 2017) (AZIMI et al., 2017) (DEEN, 2015) (HAMEURLAINE et al., 2015) (JUNG, 2017) (KOMNAKOS et al., 2008) (MACHADO et al., 2013) (MATTHEW-MAICH et al., 2016) (MILEO et al., 2010) (NAVA-MUÑOZ; MORÁN, 2012) (O'DONOGHUE; HERBERT, 2012) (THOMAS et al., 2014)
Saúde móvel	(BAE et al., 2012) (BEATTIE et al., 2014) (HAMEURLAINE et al., 2015) (KOMNAKOS et al., 2008) (MATTHEW-MAICH et al., 2016) (MICHALAKIS; CARIDAKIS, 2017) (MNATSAKANYAN et al., 2012) (O'DONOGHUE; HERBERT, 2012) (SANNINO; De Pietro, 2014) (SONG et al., 2012) (TREBATOSKI et al., 2010)
Lar de cuidados	(NAVA-MUÑOZ; MORÁN, 2012)
Autocuidado assistido	(BEATTIE et al., 2014) (MITCHELL et al., 2011)
Tele saúde	(BEATTIE et al., 2014) (CHIANG; LIANG, 2015) (THOMAS et al., 2014)

Fonte: Elaborado pelo autor

A primeira linha da tabela 7, cuidados domiciliares, refere-se ao cenário definido pela as-

sistência médica domiciliar. Nesse sentido, vários estudos no corpus deste artigo apresentam soluções diferentes. Mileo et al. (MILEO et al., 2010) propõe o uso de um sistema inteligente de saúde em casa caracterizado por uma rede de sensores sem fio e um componente de raciocínio baseado em programação de conjunto de respostas (ASP). Alirezaie et al. (ALIREZAIE et al., 2017) usa uma ontologia para integrar as medidas coletadas de fontes heterogêneas para permitir a interpretação semântica dos eventos e a percepção do contexto, enquanto Hameurlaine et al. (HAMEURLAINE et al., 2015) usa uma ontologia para representar informações sobre o contexto e um modelo baseado em regras para raciocinar sobre essas informações contextuais. Jung (JUNG, 2017) apresenta uma tecnologia doméstica inteligente usando sensores vestíveis e sensores ambientais, com o objetivo de criar uma conscientização da situação e apoiar a tomada de decisões e recomendar um tratamento adequado com base em uma taxa de risco à saúde.

A segunda linha da tabela 7, saúde móvel, refere-se à capacidade dos sistemas de adaptar o monitoramento da saúde em face da mobilidade do paciente. Nesse contexto, Komnakos et al. (KOMNAKOS et al., 2008) apresenta um estudo sobre a avaliação de potencial e desempenho da tecnologia 3.5G para fornecer aplicativos de saúde eletrônicos abrangentes, considerando o desempenho da rede de sensores juntamente com a rede 3.5G, um fator crítico para esta aplicação. Por outro lado, Bae et al. (BAE et al., 2012), levando em consideração os efeitos negativos à saúde de fatores ambientais, como poluição do ar e umidade, propõem o monitoramento de trajetórias individuais de movimento e condições ambientais, para identificar relações significativas entre esses dados como forma de apoiar os sistemas públicos de saúde. Sannino e Pietro (SANNINO; De Pietro, 2014) e Komnakos et al. (KOMNAKOS et al., 2008) salienta que o conhecimento do contexto do paciente é essencial para os serviços de aplicativos em ambientes e saúde móvel. Os primeiros propõem, para esse fim, um sistema de apoio à decisão baseado em regras; os últimos propõem uma ontologia escalável para modelagem e raciocínio.

Para a terceira linha da tabela 7, de acordo com Nava-Muñoz e Morán (NAVA-MUÑOZ; MORÁN, 2012), existem dois problemas significativos no atendimento a idosos em um lar de idosos, levando em consideração a criticidade desse tipo de ambiente. Primeiro, há falta de conscientização sobre situações que envolvem o cuidado ao idoso e, segundo, falta de informação sobre a disponibilidade e atividades de outros cuidadores para apoiar o processo de coordenação. Para lidar com esse tipo de situação, eles propõem um modelo para o design de notificações sensíveis ao contexto em ambientes críticos, tendo como principal característica considerar três fontes de contexto (o ambiente, o remetente e o destinatário da notificação).

Em relação ao autocuidado assistido e à tele saúde, a quarta e quinta linhas da Tabela 7, respectivamente, referem-se ao autocuidado assistido por meio da aplicação de tecnologias digitais. Em Beatti et al. (BEATTIE et al., 2014), com relação às pessoas que sofrem de doença pulmonar obstrutiva crônica, salienta-se que o monitoramento contínuo da saúde, comportamento e informações contextuais de um paciente fornece a capacidade de detectar um declínio em sua saúde antes que ocorra um problema. Eles propõem uma ferramenta de autogestão com reconhecimento de contexto com um módulo de previsão para gerar avisos de intervenção

apropriados, informando ao paciente que seu estado de saúde está em declínio. Chiang e Liang (CHIANG; LIANG, 2015) e Mitchell et al. (MITCHELL et al., 2011) propõe um sistema de atendimento domiciliar que armazena os contextos necessários de conhecimento em ontologias, incluindo as informações fisiológicas e as informações ambientais do paciente, fornecendo uma interface de consulta unificada com os dados contextuais em dispositivos móveis e fornecendo feedback interativo ao usuário.

3.3.4 Quais são as técnicas de modelagem de contexto empregadas na área da saúde?

Para responder à pergunta RQ4, discutimos as técnicas de modelagem de contexto mais populares. Cada um dos métodos a seguir tem seus pontos fortes e fracos, e as implementações reais dessas técnicas podem variar, dependendo do domínio do aplicativo.

De acordo com Mshali et al. (MSHALI et al., 2018), a técnica de modelagem de valor-chave modela as informações como pares de valor-chave em diferentes formatos, mas não é escalável e não é adequada para armazenar estruturas de dados complexas. Outra técnica, usando um esquema de marcação, define estruturas hierárquicas de dados usando uma linguagem de marcação, como XML. Tags de marcação são usadas para representar o formato dos dados. A modelagem gráfica consiste em uma representação diagramática de dados contextuais em nível do design, usando modelos apropriados, como a UML e outros. A modelagem orientada a objetos usa o conceito de objetos, com hierarquias de classe e relacionamento, emprega encapsulamento, reutilização e herança para representar dados de contexto. Na modelagem baseada em lógica, o contexto é representado como um conjunto de fatos, expressões e regras lógicas. A modelagem baseada em ontologia representa conhecimento e informações contextuais usando tecnologias semânticas. Estão disponíveis diferentes recursos de raciocínio e padronização (RDF¹¹, OWL¹²).

Mileo et al. (MILEO et al., 2010) apresenta um sistema doméstico inteligente com um componente de raciocínio baseado no conjunto de respostas (ASP¹³), responsável pela contínua contextualização do estado físico, mental e social do paciente. O componente de raciocínio aplica regras lógicas expressivas que visam a interpretação correta de informações contextuais incompletas ou inconsistentes. Chiang e Liang (CHIANG; LIANG, 2015) e Kuijs, Rosencrantz e Reich (KUIJS; ROSENCRANTZ; REICH, 2015) usam ontologias para armazenar contextos, informações ambientais e informações sobre o paciente. Alirezaie et al. (ALIREZAIE et al., 2017) e Hameurlaine et al. (HAMEURLAINE et al., 2015) determina o potencial de ambientes domésticos inteligentes e propõe o uso de ontologias escalonáveis para modelagem e raciocínio, não apenas sobre as medições de saúde dos pacientes, mas também sobre todas as informações contextuais, de modo a fornecer serviços de saúde adequados em residências inteligentes.

¹¹utilizada como um método para a descrição conceitual ou de modelagem de informação

¹²linguagem para definir e instanciar ontologias na Web

¹³Answer set programming é uma forma de programação declarativa orientada a problemas, baseado na semântica de modelo estável da programação em lógica

Mcheick (MCHEICK, 2014) afirma que, para modelar o contexto de um aplicativo, é preciso primeiro procurar os diferentes elementos que afetam o aplicativo, a fim de inferir elementos contextuais. Assim, a construção do contexto é dividida em dois tipos de recursos: os dinâmicos (dados médicos e ambientais) que mudam frequentemente e os elementos estáticos (nome, idade, condição médica). É proposto um formulário híbrido para modelagem de contexto, que é baseado em objetos para os componentes dinâmicos, que usa XML (um esquema de marcação para os elementos estáticos).

3.3.5 Quais são os métodos de raciocínio e fusão de dados usados em informações de ciência de contexto em saúde?

Na questão RQ5, apresentamos e discutimos as técnicas de fusão usadas para extrair dados significativos de diferentes sensores da IoT, bem como os métodos de raciocínio contextual que aparecem nos artigos do corpus. As características das informações contextuais dificultam a obtenção de um significado preciso: são necessárias quantidades substanciais de dados para entender as intenções e situações do usuário. Além disso, os serviços médicos exigem precisão. Para esse fim, dados de várias fontes e sensores são combinados para obter um resultado mais confiável, preciso e completo. Esses dados são usados para suportar componentes de raciocínio. Usando informações de contexto, o comportamento do aplicativo pode ser personalizado para uma situação específica (BALOCH; SHAIKH; UNAR, 2018; HAMEURLAINE et al., 2015). Os métodos são apresentados na Tabela 8 e são baseados na análise dos estudos que fazem parte do corpus. Frequentemente, o mesmo estudo emprega mais de um método, como pode ser visto na tabela.

De acordo com Alshurafa et al. (ALSHURAFa et al., 2014), a detecção da atividade humana independentemente de sua intensidade é essencial em muitas aplicações, especialmente no cálculo de taxas metabólicas equivalentes e na extração da ciência de contexto humano. Para esse propósito, eles usam o cluster exclusivo de k -means e um algoritmo probabilístico de cluster baseado em um modelo de mistura gaussiano. Alirezaie et al. (ALIREZAIE et al., 2017) se concentra no uso de ontologias para integrar as medidas coletadas de fontes heterogêneas, a fim de permitir uma interpretação semântica dos eventos e conscientização do contexto. O componente de raciocínio usa uma solução baseada na programação do conjunto de respostas. Da mesma forma, Hameurlaine et al. (HAMEURLAINE et al., 2015) propõe uma ontologia escalável para a integração de dados de diferentes origens, bem como um componente de raciocínio baseado em regras. Sannino e De Pietro (SANNINO; De Pietro, 2014) também propõem um sistema móvel inteligente que reconhece automaticamente o contexto, analisando dados de vários sensores e escolhendo automaticamente, através de um sistema de suporte à decisão baseado em regras, a ação a ser executada.

Chiang e Liang (CHIANG; LIANG, 2015) desenvolvem um sistema de saúde baseado em ontologia para armazenar contextos, com um mecanismo de inferência difuso para a tomada de

Tabela 8 – Métodos de fusão e raciocínio

	Método	Artigo
Fusão de dados	Estatístico	(ELHELW et al., 2009) (JUNG, 2017) (MSHALI; LEMLOUMA; MAGONI, 2018)
	Probabilístico	(BAE et al., 2012) (BEATTIE et al., 2014) (DEEN, 2015) (ELHELW et al., 2009) (MNATSAKANYAN et al., 2012) (NAVA-MUÑOZ; MORÁN, 2012) (SONG et al., 2012)
Raciocínio de contexto	Aprendizagem supervisionada	(FORKAN et al., 2015) (MITCHELL et al., 2011)
	Aprendizagem não supervisionada	(ALSHURAFUFA et al., 2014)
	Baseado em regras	(AL-NEYADI; ABAWAJY, 2009) (BEATTIE et al., 2014) (DOBRESCU, 2015) (FORKAN et al., 2015) (HAMEURLAINE et al., 2015) (MILEO et al., 2010) (SANNINO; De Pietro, 2014) (SANTOS et al., 2016)
	Lógica probabilística	(BEATTIE et al., 2014) (ELHELW et al., 2009)
	Baseado em ontologia	(ALIREZAIE et al., 2017) (KUIJS; ROSENCRANTZ; REICH, 2015) (ZERKOUK et al., 2014)
	Lógica difusa	(CHIANG; LIANG, 2015)

Fonte: Elaborado pelo autor

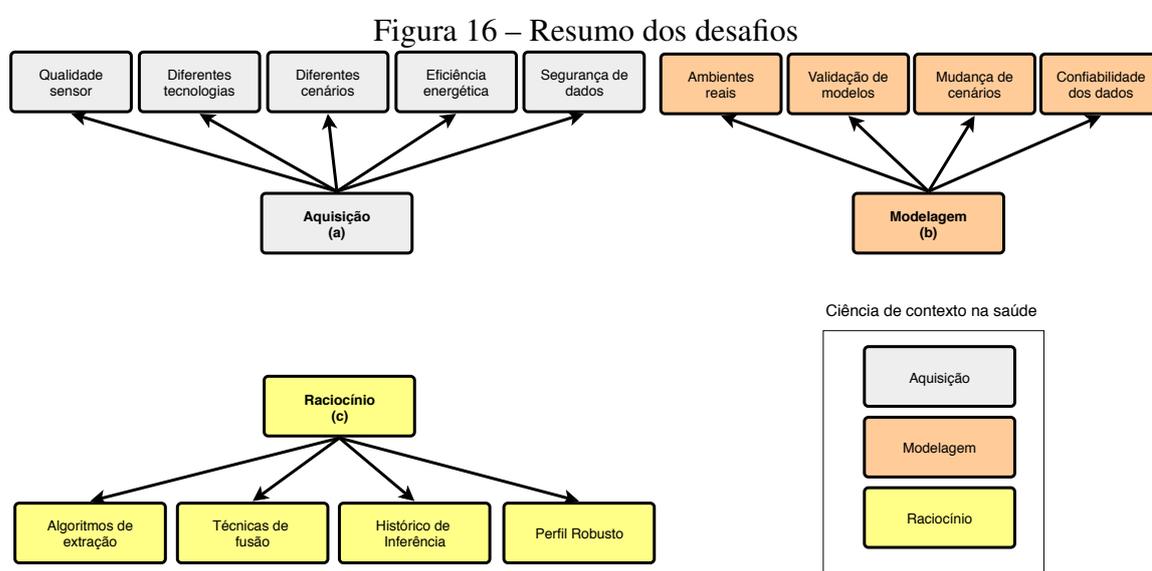
decisões. Song et al. (SONG et al., 2012) propõe um middleware sensível ao contexto para serviços móveis de saúde, especificamente para medicina preventiva. O middleware coleta as informações através do reconhecimento de diversos dados ambientais que afetam a saúde dos pacientes.

A probabilidade bayesiana também foi aplicada ao cálculo de valores de frequência repetitivos nas informações ambientais e pessoais do usuário para determinar as relações entre as informações de contexto e as práticas de medicina preventiva. Por exemplo, Forkan et al. (FORKAN et al., 2015) apresenta uma metodologia de aprendizado em duas etapas, que correlaciona atributos contextuais com valores-limite de sinais vitais, gerando no final um conjunto de regras de associação específicas para um determinado paciente: um sistema de aprendizado supervisionado é aplicado ao conjunto de dados gerado no primeiro passo, melhorando assim a precisão de prever a situação do paciente. Mitchell et al. (MITCHELL et al., 2011) apresenta uma estrutura para consultas no dispositivo móvel com um componente de aprendizado supervisionado.

3.4 Desafios e direções futuras

Após responder a todas as perguntas da pesquisa, podemos identificar os desafios inerentes ao uso de dados de contexto na área da saúde. A partir da taxonomia, destacamos os desafios

na Figura 16. Os primeiros desafios são (a) aquisição de contexto, incluindo sensores e cenários de aplicação. (GOMES et al., 2017) trata de sensores quanto à sua qualidade e melhoria dos dados de contexto fornecidos. (KOMNAKOS et al., 2008) fornece uma comparação entre várias tecnologias de sensores de baixa e alta taxa (por exemplo, Zigbee x Bluetooth) em diferentes cenários e com diferentes quantidades, incluindo o movimento do paciente, como uma maneira de tornar uma plataforma de atendimento domiciliar mais eficaz. Outros desafios para a aquisição de dados incluem eficiência energética, capacidade de resposta e robustez. Como os sensores costumam relatar seus dados para um nó central, é necessário protocolos seguros. A padronização e integração de sensores é desejável e se apresenta como outro desafio (MICHALAKIS; CARIDAKIS, 2017; DEEN, 2015), além de estender os modelos existentes a diferentes cenários, como forma de validação.



Fonte: Elaborado pelo autor

Com relação à modelagem do contexto (b), a confiabilidade e a integridade dos dados fornecidos são essenciais, sem espaço para lacunas no fornecimento de dados. Caso contrário, haverá uma fraca representação da consciência de contexto (GOMES et al., 2017). Por (CHIANG; LIANG, 2015), especialistas no domínio poderiam fornecer conhecimento mais preciso, para enriquecer as ontologias de conhecimento do contexto do usuário. A troca de cenários realizados pelo paciente pode causar múltiplas dificuldades na construção do novo contexto e sua representação de acordo com (FORKAN et al., 2015). Assim como a validação de modelos de representação contextual deve ser estendida com dados adicionais, para melhor representar a realidade (ZERKOUK et al., 2014; ALSHURFAFA et al., 2014), a expansão de protótipos para ambientes físicos reais é apontada por (MACHADO et al., 2013) como algo a ser alcançado.

Da perspectiva de (c) raciocínio, que envolve raciocínio de contexto e métodos de fusão de dados, surgem alguns desafios. Isso inclui o desenvolvimento de algoritmos para a extração e combinação de informações visuais e técnicas aprimoradas de fusão de dados para detecção

mais precisa das atividades, evitando alarmes falsos (ELHELW et al., 2009; ALIREZAIE et al., 2017). A representação gráfica das dependências e os resultados das tarefas de raciocínio podem ser aprimorados através de métodos automáticos aplicados à análise do histórico das inferências, conforme (MILEO et al., 2010). Métodos mais robustos de criação de perfil de comportamento autônomo, que melhorarão a detecção de atividades, minimizariam a carga de trabalho dos cuidadores, possibilitando o monitoramento a longo prazo dos idosos (ELHELW et al., 2009; NAVA-MUÑOZ; MORÁN, 2012).

Para concluir, podemos citar os desafios e questões de segurança e privacidade, no que diz respeito ao uso de informações de contexto em saúde. Esses problemas afetam diretamente a confiança nos sistemas que usam as informações geradas pelos sensores para criar uma percepção contextual (ZERKOUK et al., 2014; KUIJS; ROSENCRANTZ; REICH, 2015). Neste ponto, podemos destacar a aceitação como um desafio importante no uso de sensores, especialmente no que se refere à confiabilidade no uso de sistemas médicos inteligentes em casa (ZIEFLE; ROCKER; HOLZINGER, 2011), ou o uso ético e moralmente justificado da tecnologia, conforme discutido em (HOLZINGER; SCHAUPP; EDER-HALBEDL, 2008). Em (HOLZINGER; SEARLE; WERNBACHER, 2011), há uma discussão sobre a aceitação de novas tecnologias à luz da engenharia de usabilidade e acessibilidade, sobre seu desenvolvimento e uso. Da mesma forma, o uso da computação em nuvem para procedimentos intensos de computação e a insegurança das pessoas em relação ao compartilhamento de seus dados médicos pessoais impõe desafios a serem superados em relação à segurança e privacidade dos dados (MICHALAKIS; CARIDAKIS, 2017). Por fim, o acesso às informações do paciente por terceiros deve ser cuidadosamente controlado, incorporando domínios confiáveis para que usuários desconhecidos que não fazem parte do domínio não tenham acesso (AL-NEYADI; ABAWAJY, 2009).

3.5 Ciência de contexto e *diabetes mellitus*

Esta seção apresenta uma revisão de literatura que contempla trabalhos dentro do conceito de ciência de contexto relacionados à Diabetes Mellitus tipo 2 (DM2) e aprendizado de máquina. A motivação é complementar os dados coletados pela Revisão Sistemática de Literatura descrita neste capítulo aos resultados obtidos nesta seção. Dessa forma, os desafios destacados puderam ser mapeados à realidade do paciente crônico de diabetes mellitus tipo 2, fornecendo assim informações relevantes para orientar a elaboração desta tese.

A string de pesquisa utilizada foi a mesma definida na Seção 3.1.3, acrescida de termos relativos à Diabetes Mellitus tipo 2, machine learning e predição, com a string de pesquisa definida conforme quadro abaixo.

String de Pesquisa

(Context aware OR Context awareness OR Situational Aware OR Situational Awareness)
AND (Health OR HealthCare) AND (Environment data) AND (Diabetes Mellitus type 2
OR DM2) AND (Machine Learning) AND (Prediction)

A execução de buscas de artigos foi realizada no Google Scholar¹⁴. Nos resultados, a seleção de artigos ocorreu conforme o critérios descritos na Seção 3.1.4 e a avaliação da qualidade dos artigos ocorreu conforme descrito na Seção 3.1.5 e, por fim, foram selecionados artigos publicados entre os anos de 2018 e 2023 em inglês e excluindo resultados de patentes e citações. A tabela 9 apresenta os artigos finais selecionados.

A partir dos desafios identificados pela revisão de literatura, podemos traçar um paralelo entre as lacunas e desafios no uso de ciência de contexto em cuidados de saúde e as necessidades de cuidados com a diabetes mellitus tipo 2 (DM2). Um dos aspectos principais de qualquer solução de saúde eletrônica é o gerenciamento da qualidade de situações urgentes (RIVERA et al., 2019). Atualmente, essas situações são acessíveis em uma ampla variedade de sensores incorporados. A heterogeneidade de tais sensores e a diversidade de necessidades do usuário exigem qualidade de serviço de gerenciamento e adaptação a diferentes situações críticas.

Considerando que a fisiopatologia do DM2 é um processo contínuo, anormalidades críticas transitórias devem ser detectadas precocemente (BRISIMI et al., 2019; DWORZYNSKI et al., 2019). As redes de sensores capazes de fornecer dados contínuos de monitoramento fisiológico (por exemplo, glicose, pressão arterial, pulso, ritmo cardíaco) (RAMAZI et al., 2019; ABAKER; SAEED, 2021) e estilo de vida (dieta, atividade física, etc) (FARUQUI et al., 2019) têm um grande potencial para detectar essas transições e acompanhar o progresso da doença. A tecnologia da Internet das Coisas pode contribuir significativamente nessa direção, permitindo a integração de mais fatores na tomada de decisões clínicas relacionadas ao gerenciamento de diabetes (ZARKOGIANNI et al., 2015).

Devido a natureza multifatorial do DM2, abordagens de modelagem multinível devem ser aplicadas para levar em consideração todos os diferentes tipos de fatores que estão fortemente associados ao início e à evolução da doença (PAN et al., 2023). Novos métodos de análise de dados, podem ser usados para capturar relacionamentos correlatos e causais entre as variáveis. Exemplos disso, são a classificação e as abordagens baseadas em regressão (Por exemplo, regressão linear e regressão logística). Tais métodos podem ser aplicados para identificar novos biomarcadores, que estão fortemente relacionados com o início e a evolução da DM2 (MARSARO et al., 2019; ZARKOGIANNI et al., 2015; LEE; WON; LEE, 2023)

¹⁴<https://scholar.google.com.br>

Tabela 9 – Corpus de artigos Diabetes Mellitus tipo 2, machine learning e Predição

Estudo	Objetivo	Editora	Tipo
Brisimi et al. (2019)	Identificar pacientes com probabilidade de serem hospitalizados devido a complicações atribuídas ao DM2.	SagePub	Journal
Dworzynski et al. (2019)	Identificação de indivíduos com risco de desenvolver comorbidades decorrentes da DM2.	Nature	Journal
Faruqui et al. (2019)	Prever dinamicamente os níveis diários de glicose em pacientes com DM2 com base em dados de dieta, atividade física e peso.	Jmir	Journal
Aliberti et al. (2019)	Investiga modelos de predição treinados em sinais de glicose, para inferir valores futuros de nível de glicose em um paciente novo.	IEEE	Journal
Munna et al. (2020)	Detectar prevalência de complicações entre pacientes com DM2, utilizando um modelo de previsão baseado em classificação	Springer	Conference
Chen et al. (2021)	Classificação do diabetes mellitus tipo 2 com ou sem comprometimento cognitivo, usando conectividade funcional de alta ordem.	Wiley	Journal
Nanda et al. (2022)	A identificação precoce dos fatores de risco associados ao desenvolvimento da úlcera do pé diabético e classificação do tipo de úlcera.	Springer	Journal
Saravanakumar e Sabibullah (2022)	Predição de bem-estar em pacientes com DM2 através de classificação.	IJCSNS	Journal
Lee, Won e Lee (2023)	Desenvolver modelos de aprendizado de máquina (ML) que identifiquem depressão em pacientes com DM2.	PLOS	Journal
Pan et al. (2023)	Predição de risco para retinopatia diabética em pacientes com diabetes mellitus tipo 2 usando poucos indicadores.	Frontiers	Journal

Fonte: Elaborado pelo autor

3.6 Considerações Parciais

Este capítulo expôs investigações específicas relacionadas ao projeto proposto, proporcionando uma análise do estado atual da literatura sobre ciência de contexto aplicada à saúde. A

revisão de literatura delineada neste capítulo desempenha um papel fundamental na identificação dos desafios associados à implementação de tecnologias no monitoramento de pacientes portadores de doenças crônicas, como é o caso do diabetes mellitus. Inicialmente, destaca-se a problemática inerente à definição da ciência de contexto aplicada à saúde, abordando aspectos como os contextos de saúde nos quais as tecnologias são mais prevalentes, a aceitação do uso dessas tecnologias por profissionais de saúde e pacientes, além dos métodos específicos de aprendizado de máquina empregados. Por fim, visando orientar a pesquisa, uma segunda revisão foi conduzida, focalizando na perspectiva do diabetes mellitus, direcionando, assim, o trabalho para as particularidades de uma doença crônica de longa duração.

4 MODELO AIDA

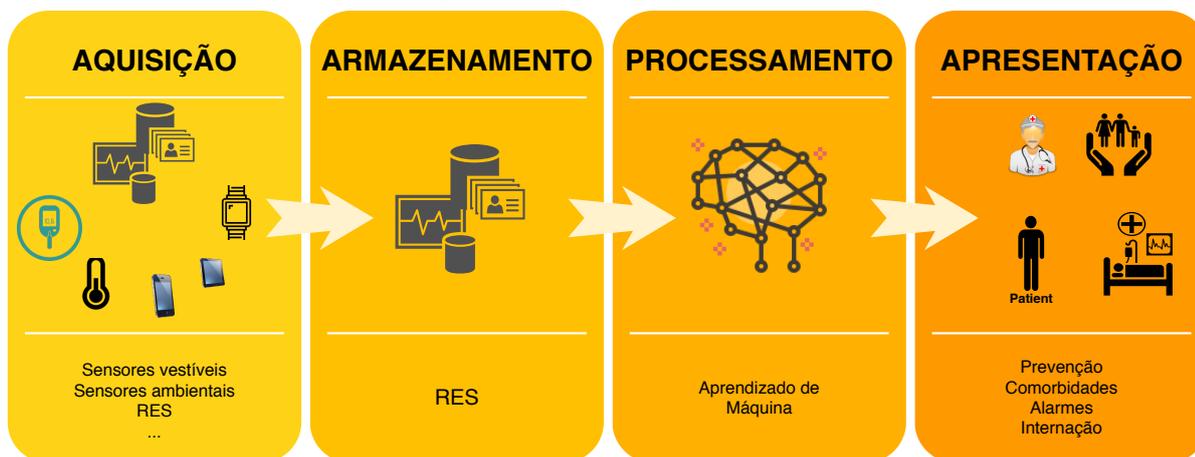
O modelo apresentado neste capítulo alia a medicina de precisão e personalizada no acompanhamento do paciente com diabetes mellitus tipo 2 (DM2). A literatura atual, conforme mencionado no Capítulo 3 e na Seção 3.5, apresenta um conjunto de oportunidades e lacunas de pesquisa no âmbito da ciência de contexto aplicada à saúde e ao diabetes. Nesse contexto, o presente capítulo expõe uma proposta de modelo inteligente fornecendo mecanismos para o desenvolvimento da medicina de precisão e personalizada no acompanhamento do paciente com diabetes mellitus tipo 2 (DM2). Conforme previamente destacado, a digitalização do prontuário eletrônico (RES) gerando conjuntos massivos de dados (LIANG et al., 2021), aliada ao uso de sensores vestíveis, para aquisição de dados no contexto do paciente (ROSSETTI; TIASE, 2022), detém o potencial para um monitoramento proativo e, em certos casos, a possível antecipação de hospitalização. A proposta central desta tese reside na contínua fusão dos dados do RES com informações obtidas diretamente dos pacientes, proporcionando suporte em tempo real para tomadas de decisões e, simultaneamente, facilitando a integração fluida entre o monitoramento do paciente e a tríade de atenção primária, profissionais médicos e equipe hospitalar. Essa integração encontra-se encapsulada na arquitetura denominada AIDA (Acrônimo com origem na frase: "uma Arquitetura Inteligente para o monitoramento de DiAbetes mellitus"). As seções subsequentes deste capítulo discorrem sobre as partes que compõem o modelo AIDA, começando com as decisões que nortearam o desenvolvimento do modelo, seguido por uma visão geral englobando os *stakeholders* inerentes ao modelo, após um detalhamento das funcionalidades e tecnologias que suportam a proposta e, concluindo com as considerações finais da abordagem.

4.1 Decisões de projeto

Nesta Seção, apresentamos as principais decisões e limitações presentes em nosso projeto. O modelo apoiará o tratamento e acompanhamento de doentes crônicos portadores de diabetes mellitus tipo 2. Inicialmente, será realizado o treinamento de um modelo de classificação e outro de regressão, utilizando dados de RES provenientes de um centro clínico do Brasil. Posteriormente, conduziremos o monitoramento contínuo dos pacientes empregando sensores. Os dados coletados desses sensores serão, por sua vez, empregados para efetuar a atualização interativa dos modelos de classificação e regressão. Essa atualização se apresenta como um meio para personalizar o tratamento e supervisão, adequadamente à singularidade de cada indivíduo. Em um contexto onde a integração de informações advindas de fontes heterogêneas, capacidade de subsídio a decisões em tempo real, e potencial de otimização da nossa aptidão para repercutir positivamente no manejo de doenças crônicas são demandas prementes e tangíveis. Nessa perspectiva, os métodos de aprendizado de máquina que usam dados observacionais rotineiros do paciente são úteis para avaliar o andamento do tratamento e informar os *stakeholders*

envolvidos em tomadas de decisão (YOU et al., 2019). Neste sentido, a Figura 17 ilustra as tecnologias e suas relações, envolvidas na construção do AIDA. A representação foi feita na forma de etapas, baseado na proposta discutida em (COSTA et al., 2018) e na taxonomia apresentada em (GUBERT; COSTA; RIGHI, 2020). A seguir explanaremos cada etapa.

Figura 17 – Tecnologias AIDA



Fonte: Elaborado pelo Autor

A primeira etapa está relacionada a aquisição dos dados. Uma multiplicidade de tecnologias pode estar envolvida na geração ou coleta desses dados, incluindo smartphones, aplicativos móveis, sensores vestíveis, entre outros. No contexto do paciente diabético, podemos citar sensores vestíveis para avaliar atividades (MSHALI et al., 2018), uso do acelerômetro e sensor de GPS do smartphone para encorajar um estilo de vida mais ativo (ALSHURFAFA et al., 2014), bem como sensores ambientais, para verificar situações de risco (SANNINO; De Pietro, 2014).

A recente abundância de dados disponibilizados em Registros Eletrônicos de Saúde (RES) proporciona uma oportunidade sem precedentes de aplicar análises preditivas para melhorar o gerenciamento de DM2 (LIU et al., 2018). Além disso, os dados gerados por sensores dos mais diversos tipos, que podem monitorar continuamente o paciente (YIN et al., 2019), servem para atualizar o estado de saúde do paciente, sendo armazenados no seu próprio RES. Assim, o RES irá consolidar as informações obtidas do paciente, para permitir consulta a qualquer momento.

Na etapa de processamento está a inteligência artificial. Neste ponto, o modelo AIDA propõe o uso da classificação supervisionada através do uso de técnicas de aprendizado de máquina e do uso de aprendizagem de máquina online para manter o modelo inteligente atualizado, para acompanhamento e personalização do tratamento do paciente com DM2. A combinação de sensores e aprendizagem de máquina traz a possibilidade de rastreamento generalizado de condições de saúde e diagnóstico de doenças em um contexto diário, extraindo *insights* médicos de sinais fisiológicos relacionados à saúde (YIN et al., 2019).

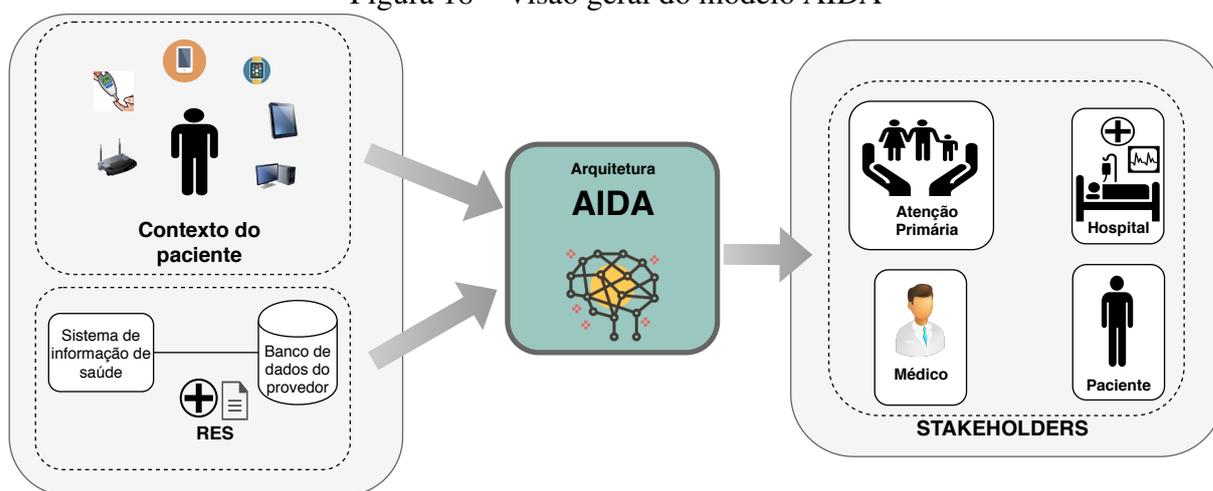
A última etapa é a de apresentação, onde são disponibilizados os resultados alcançados ao processar o conjunto de dados fornecidos pelo RES e pelos sensores, após o processamento pela IA. Neste ponto a previsão precoce do surgimento de comorbidades relacionadas ou a in-

dicação de manutenção da saúde do paciente é esperada por parte do modelo. Desta forma, as partes envolvidas no tratamento e acompanhamento do paciente diabético poderão ser atualizadas da situação, poderão receber alertas quando há deterioração da condição de saúde, previsão de melhoria ou piora do estado geral de saúde, além de verificar parâmetros de marcadores relacionados (Sociedade Brasileira de Diabetes, 2019).

4.2 Visão geral

O modelo AIDA é baseado em informações coletadas do RES e em dados coletados de sensores no paciente, para fornecer um tratamento individualizado. A Figura 18 apresenta a visão geral que foi idealizada. O conjunto de dados armazenado nos sistemas de suporte ao RES contém informações sobre todos os pacientes de um centro médico. Esse conjunto de dados será utilizado para gerar um modelo de classificação e um modelo de regressão. No contexto do paciente diabético, os dados serão coletados através de sensores médicos vestíveis (p. ex., de um smartwatch) e sensores ambientais (p. ex. de um smartphone), da mesma forma, dados podem ser inseridos pelos diversos *stakeholders* envolvidos no processo, incluindo o próprio paciente. Assim, a arquitetura foca em utilizar os dados coletados para atualizar os modelos de classificação e regressão. Como resultado espera-se conseguir a previsão antecipada de ocorrência de comorbidades ligadas a DM2, gerando assim ações corretivas por parte dos envolvidos (pacientes, médicos, equipe de atenção primária e hospital).

Figura 18 – Visão geral do modelo AIDA



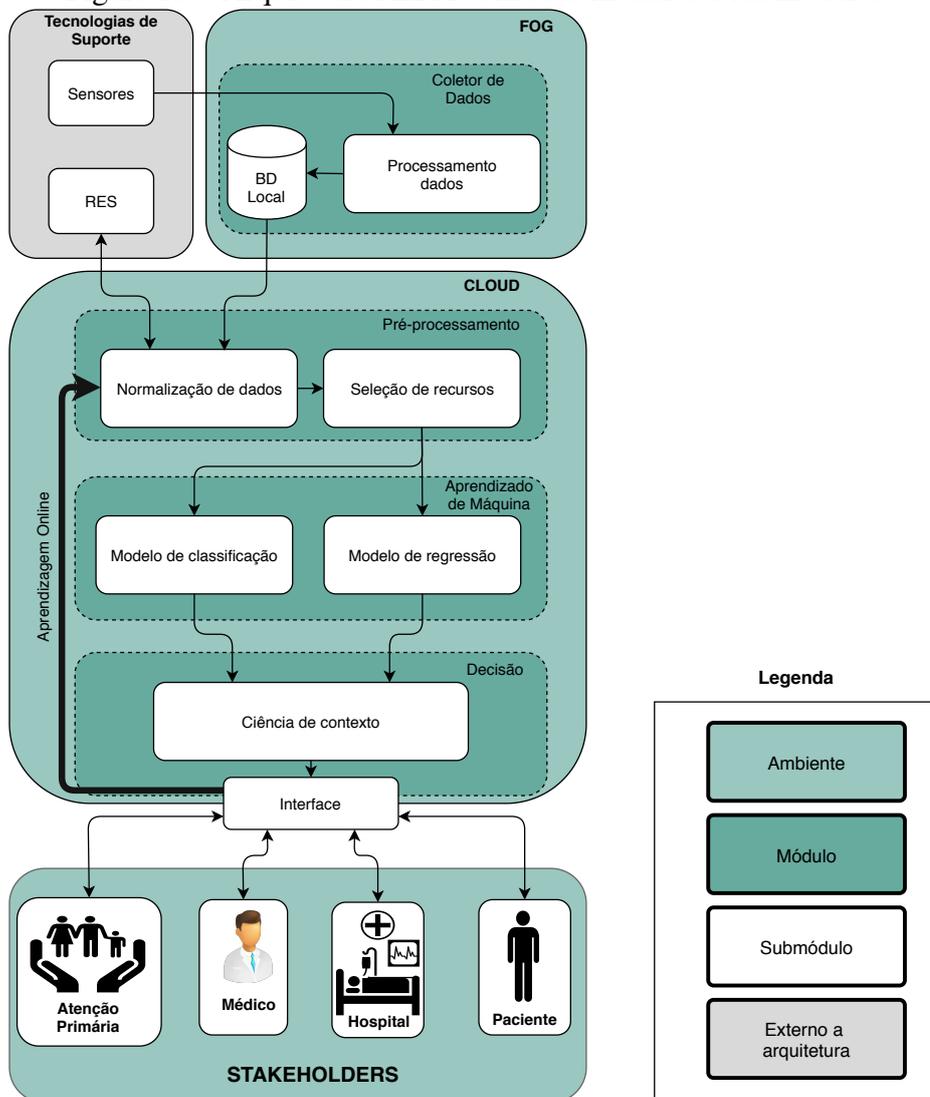
Fonte: Elaborado pelo autor

4.3 Detalhamento do modelo

O modelo AIDA é estruturado em quatro módulos fundamentais: o Módulo Coletor de Dados, o Módulo de Pré-processamento, o Módulo de Aprendizado de Máquina e o Módulo de

Decisão. O Módulo Coletor de Dados será implantado em uma infraestrutura de *Fog Computing* (Computação na Névoa), com a finalidade de estabelecer uma proximidade entre parte do processamento computacional e a arquitetura situada em torno do paciente. Esse módulo compreende dois submódulos distintos: o Submódulo de Processamento e o Submódulo do Banco de Dados Local. Por sua vez, o Módulo de Pré-processamento é composto por dois submódulos inter-relacionados: a Normalização de Dados e a Seleção de Recursos. O Módulo de Aprendizado de Máquina é constituído pelos submódulos Modelo de Classificação e Modelo de Regressão. No que tange ao Módulo de Decisão, este é composto pelos submódulos Ciência de Contexto e Interface. Importante notar que estes últimos submódulos são implementados em uma infraestrutura de nuvem computacional (Cloud). A representação visual dos módulos da arquitetura AIDA é apresentada na Figura 19, e serão detalhados nas subseções subsequentes.

Figura 19 – Arquitetura AIDA com seus módulos e submódulos



Fonte: Elaborado pelo autor

4.3.1 Módulo Tecnologias de Suporte

O módulo Tecnologias de Suporte inclui o submódulo Sensores e o submódulo RES. O submódulo Sensores serve para representar os diversos tipos de sensores em diversos contextos diferentes que podem estar presentes no monitoramento do paciente, conforme citado na Seção 3.3.3. Já o submódulo RES representa os dados legados do paciente, onde estão armazenados os dados médicos coletados em exames clínicos, pela equipe de atenção primária e de possíveis internações, além de dados vindos dos sensores conectados ao paciente.

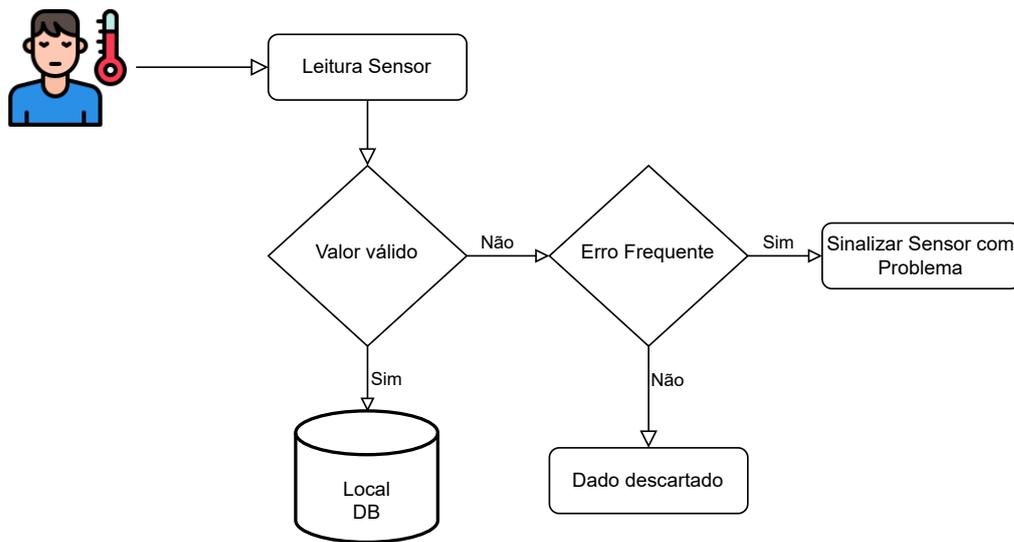
O módulo Tecnologias de Suporte está destacado em cinza na figura, pois não faz parte da descrição do modelo discutir as particularidades de implementação do conjunto de sensores no paciente e a tecnologia física de comunicação destes com o módulo Coletor de dados. Na Seção 3.3.2 os desafios do uso de sensores são discutidos. Da mesma forma, a arquitetura de funcionamento e de comunicação do RES e suas limitações não são contemplados neste modelo, mas encontram discussões com relação a interoperabilidade e não incorporação de dados sobre bem-estar, atividades físicas e hábitos alimentares no trabalho de (ROEHRS; COSTA; da Rosa Righi, 2017).

4.3.2 Módulo Coletor de dados

O módulo Coletor de dados inclui os submódulos BD Local e Processamento de dados. A função do submódulo Processamento de dados é realizar um primeiro pré-processamento dos dados das leituras que são recebidos dos sensores. Conforme os desafios discutidos na Seção 3.3.2, este pré-processamento se faz necessário para corrigir possíveis falhas de leitura dos sensores, e filtragem de valores lidos fora de parâmetros viáveis para sinais vitais. A Figura 20 apresenta um exemplo de algoritmo simples para o módulo Processamento de dados corrigir estas falhas e evitar o armazenamento de dados fora de valores viáveis. Dados anômalos com origem nos sensores podem ser causados por vários fatores, como falhas de hardware, sensor corrompido, calibração incorreta, interferência eletromagnética, entre outros (Harun Al Rasyid et al., 2019), no caso de erros de sensores e leituras erradas repetidas vezes, é gerada uma sinalização para eventual correção/troca de sensor. No exemplo da Figura 20), os valor de temperatura lido pelo sensor pode ser um valor não viável, como uma temperatura do corpo impossível de ocorrer (p. ex. 300°C) ou um valor de erro de transmissão ou leitura (p. ex. 3Ab4). Todos os dados lidos pelos sensores passarão por filtros de valores clínicos não viáveis, conforme definidos na literatura médica.

O submódulo BD Local tem o papel de armazenar temporariamente os dados coletados dos sensores. Este módulo armazena os dados antes de serem enviados para o submódulo Normalização de dados, detalhado na Seção 4.3.3, e posterior armazenamento definitivo no RES. Esse armazenamento temporário é necessário devido ao arranjo da infraestrutura física da arquitetura, visando a eficiência, persistência dos dados (KOLODZIEJ; GONZÁLEZ-VÉLEZ;

Figura 20 – Algoritmo para filtrar valores não viáveis



Fonte: Elaborado pelo autor

KARATZA, 2017) e coleta de dados em diferentes cenários em que o paciente pode se encontrar em determinado momento, e que foram discutidos na Seção 3.3.3.

4.3.3 Módulo Pré-processamento

O módulo Pré-processamento de dados inclui os submódulos Normalização de dados e Seleção de recursos. No submódulo Normalização de dados ocorrem as correções nos dados, que consiste em inserir valores ausentes e adequar *outliers*. Esta análise de dados deve ser realizada, pois em conjuntos de dados com origem em tempo real não são incomuns erros de leitura e de processamento. Como alguns fatores dependem de características do paciente (como a idade ou o gênero), essas devem ser levadas em consideração, pois podem influenciar significativamente os intervalos válidos de valores de sinais vitais e de resultados de exames laboratoriais.

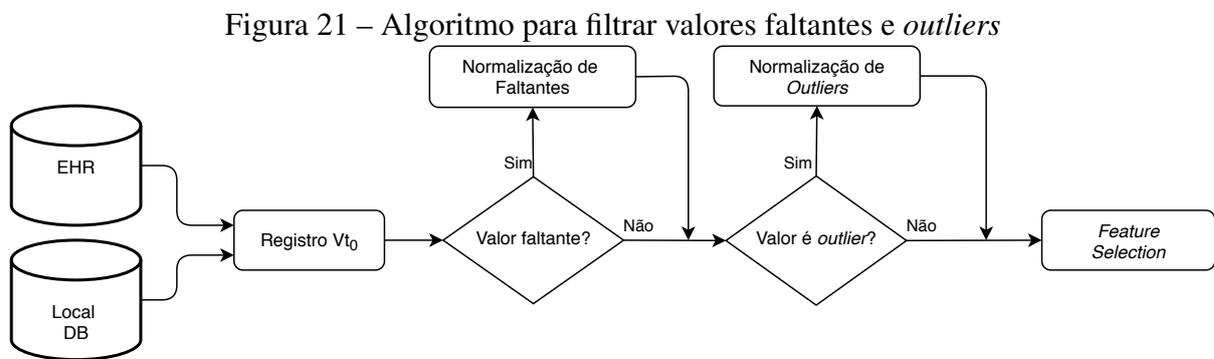
Um desafio significativo para o aprendizado de máquina é a ausência de itens de dados para um determinado paciente. Uma alternativa que pode ser considerada para lidar com dados ausentes é a imputação de dados, pois uma proporção substancial deles podem acontecer em um conjunto de dados. A imputação de valores ausentes pode ser considerada uma das maiores tarefas do pré-processamento, pois a maioria dos conjuntos de dados médicos geralmente está incompleta, sendo que simplesmente remover os casos incompletos dos conjuntos de dados originais pode trazer mais problemas do que soluções. Uma das técnicas mais utilizadas para inserir dados ausentes é a imputação média (MANIRUZZAMAN et al., 2018), que pressupõem que a média de uma variável acaba sendo a melhor estimativa e consiste na inserção da média

do valor da variável no local do dado faltante, conforme a função 4.1:

$$Vt_0 = avg(Vt_1 + \dots + Vt_n) \quad (4.1)$$

onde Vt_0 é o registro faltante, e $avg(Vt_1 + \dots + Vt_n)$ é a média dos registros presentes. Além da imputação média, podem ser utilizadas outras técnicas de imputação de dados faltantes, como a maximização regular de expectativa, imputação múltipla e imputação kNN (Duy Le; BEURAN; TAN, 2018). Desta forma, um método adequado para imputação de valores faltantes auxilia a produzir um conjunto de dados de boa qualidade.

O tratamento de *outliers* também é uma questão de fundamental importância na etapa de pré-processamento, já que dados fora de padrão podem levar a classificações com baixa precisão. Desta forma, o tratamento de valores pode ser feita usando a imputação baseada na média de dados (conforme visto na substituição de valores ausentes). Uma diferença importante entre os dois tratamentos (dados faltantes e *outliers*), é que a detecção de *outliers* é mais complexa, visto que o dado está no conjunto, mas se apresenta de forma discrepante com relação aos demais dados. Uma solução para detectar este dado fora de padrão é o uso da técnica de intervalos inter-quartil (IQR) proposta por (MANIRUZZAMAN et al., 2018). A Figura 21 exemplifica um algoritmo para o tratamento de dados ausentes e para substituição de *outliers*.



Fonte: Elaborado pelo autor

Outra função do submódulo Normalização de dados é colocar os valores dos recursos em um formato normalizado. A normalização é uma técnica aplicada como parte da preparação de dados para o aprendizado de máquina. O objetivo da normalização é alterar os valores do conjunto de dados para usar uma escala comum, sem distorcer diferenças nos intervalos de valores (CHRISTOBEL; SIVAPRAKASAM, 2013). Em um conjunto de dados pode haver uma grande diferença na escala dos números, o que pode causar problemas quando se tenta combinar os valores como recursos durante a modelagem, assim as magnitudes dos valores são escalonadas para valores consideravelmente baixos (ACHARYA et al., 2017). Desta forma, o modelo AIDA irá aplicar a normalização *Z-score* como uma das etapas de pré-processamento, conforme a equação:

$$Z = \frac{x - \mu}{DP_{pop}} \quad (4.2)$$

onde: μ é a média da população, x é um valor dos dados e DP_{pop} é o desvio padrão da população.

Por fim, dentro do módulo Pré-processamento, há o submódulo Seleção de recursos, onde acontece a definição de recursos relevantes a partir do conjunto de dados original. A seleção de recursos é a técnica de pré-processamento que seleciona o subconjunto ideal a partir de todos os recursos disponíveis (ALPAYDIN, 2016). Devido a natureza dos dados médicos, eles podem conter dados irrelevantes e redundantes, reduzindo assim a precisão da previsão. Desta forma, selecionar os recursos mais relevantes, contribui para a importância do pré-processamento na qualidade da classificação final, aumentando a eficiência da aprendizagem, a precisão preditiva e reduzindo a complexidade dos resultados aprendidos (GANDHI; PRAJAPATI, 2014).

O modelo AIDA fará a seleção de *features* utilizando o método *F-score*, que é um método simples para seleção de características. Este método é dado pela comparação direta das distâncias entre as médias de duas distribuições, em relação às suas variâncias (HACKELING, 2014). A métrica *F-Score* é calculada por:

$$f_i \equiv \frac{(\bar{m}_i^{C_1} - \bar{m}_i) + (\bar{m}_i^{C_2} - \bar{m}_i)}{\sigma_i^{C_1} + \sigma_i^{C_2}} \quad (4.3)$$

Sendo que a característica i que maximiza f_i é referente à melhor separação entre duas distribuições C_1 e C_2 , portanto, essa métrica pode ser utilizada como critério para seleção das características mais relevantes a serem aplicadas à um classificador.

4.3.4 Módulo Aprendizado de máquina

Neste módulo são utilizadas técnicas de aprendizado de máquina responsáveis por prover a capacidade de detecção e predição inteligente de crises ao modelo AIDA. O módulo Aprendizado de máquina inclui os submódulos Modelo de classificação e Modelo de regressão. As principais tarefas envolvidas neste módulo são comuns as técnicas de aprendizado de máquina: treinamento, teste, validação e classificação/regressão. O treinamento, teste e validação serão descritos no Capítulo 5.

No submódulo Modelo de classificação, os dados históricos de pacientes armazenados no RES e originários dos sensores, serão utilizados como dados para treinamento. O objetivo deste submódulo é uma classificação binária, do tipo 0 ou 1, ou seja, determinar se o paciente está em crise ou não, conforme mostrado na Figura 22. Várias técnicas de aprendizado de máquina tem sido utilizadas para classificação em dados de diabetes, entre elas podemos citar SVM (*support vector machine*) e classificação bayesiana (ABHARI et al., 2019), Para além destas, redes neurais recorrentes e LSTM (long short-term memory) também tem sido utilizadas especificamente para área de saúde, devido as características de intervalos e atrasos relativamente

longos em séries temporais (JUNWEI et al., 2019). Adicionalmente, algoritmos como regressão logística, árvores de decisão, floresta aleatória, *gradient boosting*, *XGBoost* e *LightGBM* são relevantes na classificação de pacientes com diabetes mellitus tipo 2, cada um com suas respectivas capacidades, desde interpretabilidade até otimização de precisão, contribuindo para insights abrangentes e soluções eficazes no entendimento e gerenciamento dessa condição médica complexa (ABHARI et al., 2019; RENGASAMY; ROTHWELL; FIGUEREDO, 2021).

Figura 22 – Submódulo de classificação



Fonte: Elaborado pelo autor

Da mesma forma que no submódulo de classificação, o modelo de regressão utilizará dados históricos dos pacientes armazenados no RES e com origem nos sensores, para treinamento do modelo. O objetivo deste submódulo é uma regressão em que as saídas são contínuas, com o objetivo de prever o tempo para o paciente entrar em uma crise, baseado nos seus dados históricos. Na figura 23, podemos visualizar as entradas do modelo, representadas pela base de dados do RES e pelas informações coletadas pelos sensores, alimentando um modelo baseado em técnicas de aprendizado de máquina. A literatura apresenta diversas abordagens que são utilizadas para realizar regressão em dados históricos, como florestas aleatórias (*Random Forest*) (RODRÍGUEZ-RODRÍGUEZ et al., 2019) para prever nível de glicemia, redes neurais (ZOU et al., 2018) e LSTM (MASSARO et al., 2019) para prever o surgimento do DM2.

Figura 23 – Submódulo de regressão



Fonte: Elaborado pelo autor

Como apresentado nos submódulos de classificação e regressão, as fontes de dados são o registro eletrônico de saúde e os dados fornecidos por sensores. Algumas peculiaridades a

respeito destes dados precisam ser destacadas. Primeiro, que os dados de registro eletrônico possuem uma dimensão temporal extremamente importante, assim a mineração nestes dados pode permitir uma compreensão maior da progressão da doença e repostas ao tratamento, por outro lado estes dados apresentam intervalos de tempo heterogêneos, esparsos, de alta dimensão e irregulares (REDDY; AGGARWAL, 2015).

Em segundo lugar, ao contrário dos dados de registro eletrônico, os dados do sensor são geralmente representados como séries temporais numéricas, regularmente medidas no tempo em alta frequência. Exemplos desses dados são dados fisiológicos obtidos pelo monitoramento regular dos pacientes. Os dados do sensor para um sujeito específico são medidos por um período muito mais curto de tempo, geralmente de alguns minutos a vários dias (REDDY; AGGARWAL, 2015). Assim a aplicação de uma técnica específica de aprendizado de máquina, que consiga lidar bem com as diferentes dimensões temporais dessas duas fontes de dados (RES e sensores) e capaz de aprender dependências de longo prazo, é indicada.

4.3.5 Módulo Decisão

O módulo Decisão inclui os submódulos Ciência de contexto e Interface. O submódulo Ciência de contexto realiza a análise dos dados recebidos do módulo Aprendizado de máquina. Os valores resultantes dos modelos de classificação e regressão são analisados, como forma de verificar as condições atuais de saúde do paciente e o curso da doença. A partir desta análise, os *stakeholders* podem ser atualizados a respeito da situação de saúde do paciente, do andamento do tratamento e, como forma geral, do curso da doença.

Conforme ilustrado na Figura 22 o resultado do modelo de classificação será 0 ou 1, neste caso representando que o paciente não está em crise (0) ou está em crise (1). Por sua vez, conforme mostra a Figura 23, o resultado do modelo de regressão poderá ter como resultado um valor entre 0 e 1, o que irá indicar o tempo até o paciente atingir uma crise. A Tabela 10 mostra um exemplo de possíveis ações decorrentes dos valores resultantes do módulo Aprendizado de máquina.

A Tabela 10 sugere possíveis ações que o módulo Ciência de contexto pode executar. A sugestão da tabela é de que o resultado decorrente do módulo de regressão (Figura 23) seja combinado com o valor da glicemia para que a decisão seja tomada com mais precisão. Desta forma, o envolvimento dos *stakeholders* acontece na linha do atendimento médico padrão para os casos de acompanhamento de pacientes com DM2. Como pode ser visto na Figura 5, o fluxograma de tratamento de paciente com DM2 (DUNCAM et al., 2013) consiste em um ciclo de atendimentos com definições de metas do tratamento, intercalados com ciclos de monitoramento e reavaliações a cada 3 a 6 meses. Conforme pode ser visto em várias situações, de contextos diferentes do monitoramento do paciente com DM2, a modificação da dieta é a base do tratamento. Apesar da natureza geralmente progressiva da doença, muitos pacientes recém-diagnosticados com DM2 e que seguem os conselhos alimentares mostram melhora no controle

Tabela 10 – Tabela de possíveis decisões do submódulo Ciência de contexto, baseada nos resultados do módulo Aprendizado de máquina

Tempo (dias)	Glicemia	Stakeholder	Conduta
0	Alta	Hospital / Médico/ Paciente	Deslocamento imediato ao hospital ou serviço de saúde mais próximo
	Média	Hospital / Médico/ Paciente	Deslocamento imediato ao Hospital ou serviço de saúde mais próximo
	Baixa	Hospital / Médico / Paciente	Deslocamento imediato ao serviço de saúde mais próximo
<7	Alta	Paciente / Atenção Primária / Médico	Acompanhamento intensivo da atenção primária e médico do paciente
	Média	Paciente / Atenção Primária / Médico	Acompanhamento intensivo da atenção primária e Restrição na ingestão de carboidratos e açúcares
	Baixa	Paciente / Atenção Primária	Acompanhamento intensivo da atenção primária e Restrição na ingestão de carboidratos e açúcares
7 - 14	Alta	Paciente / Atenção Primária	Acompanhamento intensivo da atenção primária e restrição na ingestão de carboidratos e açúcares
	Média	Paciente /Atenção Primária	Acompanhamento intensivo da atenção primária e restrição na ingestão de carboidratos e açúcares
	Baixa	Paciente / Atenção Primária	Manejo na alimentação, com restrição na ingestão de carboidratos e açúcares
15 - 30	Alta	Paciente / Atenção Primária	Manejo na alimentação, com restrição na ingestão de carboidratos e açúcares, acompanhamento da Atenção Primária
	Média	Paciente / Atenção Primária	Manejo na alimentação, com restrição na ingestão de carboidratos e açúcares, acompanhamento da Atenção Primária
	Baixa	Paciente / Atenção Primária	Manejo na alimentação, com restrição na ingestão de carboidratos e açúcares
>30	Alta	Paciente	Manter hábitos saudáveis: - Alimentação - Exercícios Físicos
	Média	Paciente	Manter hábitos saudáveis: - Alimentação - Exercícios Físicos
	Baixa	Paciente	Manter hábitos saudáveis: - Alimentação - Exercícios Físicos

Fonte: Elaborado pelo autor

glicêmico a ponto de impedir ou retardar a necessidade de agentes orais e insulina. Para aqueles que necessitam de terapia medicamentosa, pode-se esperar que a atenção aos conselhos dietéticos melhore ainda mais o controle glicêmico (DEFRONZO et al., 2015)

Por fim, o submódulo Interface realiza a comunicação entre o núcleo da arquitetura e os *stakeholders*. Essa comunicação ocorre nos dois sentidos, a) da arquitetura para os *stakeholders* e b) dos *stakeholders* para a arquitetura. Da arquitetura para os *stakeholders*, a comunicação ocorre como forma de alarmes e avisos. Dependendo da avaliação realizada no módulo Ciência de contexto, com base nos dados fornecidos pelo módulo Aprendizado de máquina, podem ser notificados um ou mais *stakeholders*, inclusive o próprio paciente. De acordo com a gravidade da situação do paciente, os modelos de classificação e regressão poderão gerar valores que vão do "paciente estar em crise" a valores do tipo "o paciente entrará em crise em alguns dias". Deste modo, poderão ser notificados o próprio paciente, para que tenha uma atitude mais saudável frente ao curso da doença, serviço de atenção primária, para que realize visitas mais frequentes ao paciente, e ainda o próprio médico ou o hospital, caso o paciente necessite de uma internação ou avaliação mais detalhada do curso da doença.

Por sua vez, a comunicação dos *stakeholders* para a arquitetura ocorre com o objetivo de coletar dados atualizados do paciente, conforme pode ser visto na Figura 19. Essa coleta pode acontecer de diversas formas, no caso da equipe de atenção primária e do médico, ela pode ocorrer através da atualização do registro eletrônico de saúde (RES), com dados de sinais vitais colhidos no consultório, em exames laboratoriais, acompanhamentos domiciliares ou ainda em hospitalizações. Outra forma de coleta de informações ocorre através dos sensores, que podem estar em quartos hospitalares ou em sensores conectados ao paciente.

Os dados dos sensores serão coletados *on the fly* no paciente, como forma de manter os modelos atualizados. Para isso o módulo Interface irá receber os dados atualizados de forma contínua pelos *stakeholders* utilizando a técnica de aprendizado de máquina online. O aprendizado online é uma família de métodos de aprendizado de máquina, em que o objetivo é realizar alguma tarefa preditiva (ou qualquer tipo de tomada de decisão), aprendendo com uma sequência de instâncias de dados (ZHENG et al., 2017). O objetivo do aprendizado online é maximizar a precisão e correção da sequência de previsões, dado o conhecimento de respostas corretas para tarefas anteriores de previsão e, possivelmente, informações adicionais (HOI et al., 2018).

Desta forma, como apresentado na Figura 19, os dados coletados serão enviados ao módulo Pré-processamento para que o processamento de tratamento dos dados seja realizado e, então possam ser atualizados os modelos de classificação e regressão. Semelhante aos métodos tradicionais de aprendizado de máquina (em lote), as técnicas de aprendizado online podem ser aplicadas para resolver uma variedade de tarefas em uma ampla variedade de domínios de aplicativos no mundo real. Algoritmos de aprendizado online podem ser derivados para tarefas de aprendizado supervisionado, onde uma das tarefas mais comuns é a classificação binária. Além desta, podemos destacar a análise de regressão de séries temporais, onde os dados chegam naturalmente de forma sequencial (HOI et al., 2018).

4.4 Considerações parciais

Este capítulo descreveu as características e conceitos envolvidos na proposta de uma arquitetura inteligente para gerenciamento da diabetes. Para finalizar o capítulo, algumas considerações devem ser feitas a respeito da concepção física da arquitetura AIDA. Estas se referem a proposta de implementação de seus módulos funcionais.

Para dar suporte à arquitetura AIDA, o módulo Coletor de dados é desenvolvido em uma Fog (computação em neblina) e os módulos responsáveis pelo pré-processamento, aprendizado de máquina e decisão se localizam em uma Cloud (computação na nuvem). Para muitos aplicativos de assistência médica, uma arquitetura simplificada de sensores em nuvem pode não ser viável, quer seja por políticas hospitalares (KRAEMER et al., 2017), quer seja pelos problemas decorrentes da transmissão de dados e a alternância de contexto do paciente. Para isso, um certo nível de processamento é necessário próximo ao dado coletado pelo sensor, garantindo assim a correção de possíveis erros de leitura dos sensores (GOMES et al., 2017) e armazenamento temporário dos dados (ELHELW et al., 2009; VITABILE et al., 2019).

Da mesma forma, em muitos sistemas de monitoramento de serviços de saúde, servidores remotos na nuvem foram usados para armazenar e processar uma grande quantidade de dados coletados através de sensores (GIA et al., 2015). Assim, a arquitetura AIDA prevê o pré-processamento dos dados e aprendizado de máquina, que exigem maior capacidade computacional se localizem na nuvem. A computação em nuvem pode fornecer acesso de rede onipresente, sob demanda e conveniente a recursos de computação, que podem ser compartilhados e provisionados rapidamente, com esforço mínimo de interação (MELL; GRANCE, 2011).

5 MATERIAIS E MÉTODOS

Neste capítulo será caracterizada a origem e os dados utilizados para execução da avaliação da arquitetura AIDA. Além disso, abordamos os procedimentos e as especificidades do treinamento de aprendizado de máquina para classificação e predição. A avaliação da arquitetura se concentrou em experimentos relacionados aos mecanismos do modelo AIDA e as contribuições propostas. Buscamos seguir abordagens clássicas de avaliação para sistemas inteligentes baseados em aprendizagem de máquina supervisionada, aplicadas em doenças crônicas (DRIT-SAS; TRIGKA, 2022; HASAN et al., 2020; KOPITAR et al., 2020). Foram utilizados dados de sinais vitais, exames laboratoriais e informações de medicamentos usados pelos pacientes, para dar suporte às avaliações. A avaliação da arquitetura AIDA é realizada por profissionais de saúde envolvidos na atenção primária e por parte dos pacientes DM2. A avaliação de aceitação da arquitetura inclui um estudo baseado no TAM (Technology Acceptance Model) com avaliações qualitativas dos resultados. Da mesma forma, a avaliação por parte dos pacientes foi realizada com entrevistas semi estruturadas norteadas pelo SUS (System Usability Scale), visando compreender a importância do uso do aplicativo/arquitetura na experiência do usuário (MARTINS et al., 2015). Os experimentos realizados são mostrados na Figura 24.

Figura 24 – Fluxo dos Experimentos realizados



Fonte: Elaborado pelo autor

Desta forma, este capítulo apresenta na Seção 5.1 as informações sobre a declaração de ética do estudo. Em seguida, na Seção 5.2 os dados utilizados para implementação da arquitetura proposta, como fontes de dados, preparação dos dados e construção do conjunto dos dados para treinamento e testes com os modelos de aprendizado de máquina. Na Seção 5.3 são abordados os métodos de avaliação de aceitação de tecnologia, no que se refere aos profissionais de saúde e pacientes e por fim métodos utilizados para treinamento e teste dos dados, utilizando aprendizado de máquina.

5.1 Declaração de Ética

Nesta investigação, foi oportunizado manipular dados reais provenientes de um estabelecimento clínico de saúde localizado no Brasil. A materialização dessa colaboração se deu por meio de uma parceria entre a Universidade do Vale do Rio dos Sinos (UNISINOS) e o Centro Clínico da Universidade do Vale do Taquari (Univates) em Lajeado - RS. O mencionado centro clínico disponibilizou o conjunto de dados necessário para conduzirmos nossa pesquisa. É relevante destacar que este projeto foi devidamente submetido à avaliação e obteve aprovação dos comitês de ética da UNISINOS, identificado sob o registro número 4.585.722, bem como da Univates, sob a identificação de número 4.235.499.

5.2 Materiais

Esta seção detalha a origem dos dados que foram utilizados para construir o dataset utilizado nos testes da arquitetura AIDA. Apresentamos e caracterizamos a instituição e a população de origem dos dados. Discutimos a coleta e tratamento dos dados para construção da base de conhecimento, as técnicas de pré-processamento dos dados, as técnicas de preenchimento de dados faltantes e o treinamento de modelos de classificação e predição.

5.2.1 Caracterização Centro Clínico

O Centro Clínico¹ foi inaugurado em março de 2016 e implantado de acordo com o avanço da grade curricular do bacharelado em Medicina da Univates. Atualmente estão em funcionamento no local o atendimento em mais de 17 especialidades, médicas² além de um ambulatório de Diabetes e Hipertensão, laboratórios de exames e procedimentos ambulatoriais, como ecocardiografia, ultrassonografias, endoscopias, entre outros exames³.

A origem da população atendida compreende 49 municípios da região de Lajeado/RS, correspondente a 16ª Coordenadoria Regional de Saúde⁴. Os pacientes são encaminhados ao centro clínico por meio das prefeituras, via requisições de consulta e exames emitidos pelo SUS. No ano de 2021 foram realizadas 7.564 consultas nas especialidades oferecidas destas, 628 consultas foram no consultório de endocrinologia.

Dentro do espectro do DM2, são acompanhados rotineiramente mais de 300 pacientes pelo serviço de endocrinologia do centro clínico. Os dados destes pacientes são coletados na recepção às consultas e, em parte, durante a anamnese do paciente. Da mesma forma, dados de

¹<https://www.saudeunivates.com.br/>

²alergia e imunologia clínica, gastroenterologia, anestesiologia, geriatria, cardiologia, hematologia, dermatologia, infectologia, endocrinologia, nefrologia, neurologia, neuropediatria, oftalmologia, otorrinolaringologia, pneumologia, psiquiatria e reumatologia.

³radiografias simples, endoscopia digestiva alta e baixa, espirometria, Holter, Monitorização Ambulatorial da Pressão Arterial (Mapa), eletrocardiograma, biópsias, procedimentos dermatológicos e de otorrinolaringologia.

⁴<https://saude.rs.gov.br/16-crs-lajeado>

glicose capilar são coletados pelos pacientes de forma domiciliar, anotados em uma ficha de acompanhamento, que posteriormente são levadas à consulta e utilizadas para a detecção de casos de hipo ou hiperglicemia.

5.2.2 Dados empregados

Os dados utilizados para os experimentos de aprendizado de máquina deste trabalho são provenientes do registro eletrônico de saúde de pacientes, da base de dados do Centro Clínico da Univates. São informações demográficas e informações de acolhimento, que são compostas de sinais vitais coletados na recepção ao paciente. As evoluções, que são dados coletados e anotados pelo médico durante a consulta, envolvem resultados de exames laboratoriais, alguns sinais vitais que são checados novamente, informações sobre uso de medicamentos, informações sobre vícios, além de resultados da anamnese. Também foram coletados dados de sinais vitais no contexto do paciente, através de um aplicativo, e que serviram para os experimentos de predição.



Fonte: Elaborado pelo autor

A Figura 25 mostra como os dados foram preparados para comporem as bases utilizadas nos experimentos. A base de dados para classificação foi construída a partir dos dados coletados do RES dos pacientes, compreendendo os dados estruturados de sinais vitais e dados não estruturados das evoluções. A partir dos dados da evolução foram extraídos dados de exames laboratoriais, informações sobre vícios e correções da variável de desfecho. Neste ponto, houve a integração dos dados das variáveis preditoras com a variável de desfecho, resultando na base para os experimentos de classificação. A base de dados de predição, foi construída a partir dos dados coletados no contexto do paciente, através do aplicativo AIDA, desenvolvido para este fim e apresentado na Seção 5.2.2.4. Através do uso do aplicativo, foram coletados dados de sinais vitais e da variável de desfecho. Após houve a integração destes dados para a posterior disponibilização da base de dados que foi utilizada para experimentos de predição.

5.2.2.1 Sinais Vitais de Acolhimento

Os dados de sinais vitais dos pacientes do centro clínico, consistem em dados coletados pela equipe de enfermagem na chegada do paciente para a consulta. Neste momento, diversos dados são aferidos e digitados no RES do paciente. As coletas de alguns dados são repetidos no momento da anamnese e anotados na evolução do paciente pelo médico, durante a consulta. Na Tabela 11 podemos ver uma parte da base RES dos pacientes do centro clínico.

Tabela 11 – Fotografia temporal de dados coletados no sistema do centro clínico.

E ^a	S ^b	I ^c	PaD ^d	PaS ^e	FC ^f	FR ^g	T ^h	P ⁱ	IMC ^j	A ^k	GC ^l
...
12/04/2016	F	29.44	80.0	130.0	75.0	16.0	36,2	106	35	174.0	89
27/06/2016	M	29.65	80.0	130.0	64.0	17.0	36,3	106	35	174.0	
10/08/2016	F	29.77	70.0	120.0	75.0	18.0	36,4	105,5	34,8	174.0	
31/08/2016	F	29.82	90.0	130.0	72.0		36,6	106,85	35,3	174.0	
19/09/2016	M	29.88	80.0	120.0	76.0		36,8	106,7	33,7	178.0	
26/09/2016	M	29.9	80.0	130.0	88.0	18.0	35,6	104	32,8	178.0	
14/12/2016	M	30.11	80.0	130.0	72.0		35,9	98	32	175.0	111
...

^aEntrada (Data de Atendimento); ^bSexo; ^cIdade; ^dPressão Arterial Diastólica; ^ePressão Arterial Sistólica; ^fFrequência Cardíaca; ^gFrequência Respiratória; ^hTemperatura corporal; ⁱPeso (Kg); ^jÍndice massa corporal; ^kAltura (cm); ^lGlicose Capilar

Fonte: Elaborado pelo autor

Os dados do RES visualizados na Tabela 11, representam as informações (colunas) sob as quais serão processados os modelos de classificação. Destes dados, foram retirados as informações que poderiam identificar o paciente (neste caso, o número do atendimento e do prontuário). Além disso, podem ser visualizadas células sem nenhum dado (células em branco), que correspondem a dados que, por algum motivo, não foram coletados ou que foram coletados mas não foram armazenados no RES.

5.2.2.2 Dados de Evolução

Durante a consulta ao médico especialista, também podem ocorrer coleta de dados do paciente, muitas vezes repetindo parcialmente a coleta dos dados do acolhimento. Estes dados ficam armazenados em um campo denominado evolução. Além de sinais vitais também ficam armazenados neste campo outras informações, como por exemplo, relatos do paciente e valores de exames laboratoriais. Na tabela 12, pode ser visto o campo evolução de um paciente, com data e hora da consulta. Os campos atendimento e prontuário foram omitidos para a desidentificação do paciente.

A partir do campo evolução, foram coletados outro conjunto de dados dos pacientes. Pode-

Tabela 12 – Exemplo de dados no campo evolução, coletados no sistema do centro clínico.

ENTRADA	EVOLUÇÃO
31/8/2016 8:46	<p>#HAS #HIPERFERRITINEMIA- HEMOCROMATOSE HEREDITÁRIA GENE H63D #HETEROZIGOTO. #REALIZOU 7 SANGRIAS, ÚLTIMA 10/11/15</p> <p>Paciente relata melhora da tontura, relatando episódios somente se fica muito tempo sem se alimentar. Relata parestesia em MSE no dia da consulta pela manhã antes de vir ao Centro Clínico.</p> <p>....</p> <p>MEDICAMENTOS EM USO: METROPOLOL 100mg 1/2 2x ao dia LOSARTANA 50mg 2x/dia ALOPURINOL 100mg 3x/dia CIPROFIBRATO (parou de tomar há 2 meses por questões financeiras)</p> <p>Exames laboratoriais de 18/08/16 em anexo no GED: -Triglicerídeos: 743; 203 (27/11/15), 950 (25/11/14) -glicose: 125; 109(27/11/15) -Creatinafosfoquinase:575 -HDL:30 -Colesterol Total: 228</p> <p>HGT jejum do dia 31/08/16: 97</p> <p>Exame físico: PA:140/100 TIREOIDE: SEM MASSA E NÓDULOS PALPÁVEIS. LINFONODOS:AUSÊNCIA DE LINFOADENOPATIAS</p> <p>....</p>

Fonte: Elaborado pelo autor

mos classificar estes dados em vícios⁵, uso e tipos de medicamentos⁶ e Exames laboratoriais⁷. Os dados foram coletados por especialistas, através da análise do campo evolução. A revisão do campo evolução pelos especialistas, também serviu para a revisão do dado de hipoglicemia, que no caso deste trabalho, serviu como variável alvo para os algoritmos de classificação e predição. Sendo que o objetivo final do modelo seria classificar o paciente em crise hipoglicêmica e/ou prever se o curso da doença está levando o paciente a uma crise hipoglicêmica.

⁵Etilismo, Tabagismo

⁶Sulfonilureia, Glinida, Metformina, Pioglitazona, Inibidor de dpp4, Inibidor de SGLT2, inibidor de alfa-glicosidase, Insulina

⁷TFG, Potássio, Colesterol Total, HDL, LDL, Triglicéridos, Vitamina B12, HbA1c

5.2.2.3 Variáveis analisadas para Classificação

Após a coleta dos dados do RES dos pacientes, foi realizada uma seleção de variáveis a serem utilizadas no processo de classificação, separando as variáveis preditoras da variável de interesse, que neste caso é glicose capilar (vide Tabela 11). Também foram removidas algumas variáveis que apresentavam alta taxa de dados ausentes, como dados de frequência respiratória e dados de exames laboratoriais, conforme detalhado na Seção 5.2.2.2. Pode ser visto na Tabela 13, os dados resultantes e que foram utilizados para o modelo de classificação.

Tabela 13 – Detalhamento das variáveis utilizadas para classificação

Origem	Variável	Tipo de dado	Utilização no Algoritmo
Acolhimento	Sexo	Catégorico	Preditora
	Idade	Ponto Flutuante	Preditora
	Pressão Arterial Diastólica	Inteiro	Preditora
	Pressão Arterial Sistólica	Inteiro	Preditora
	Frequência Cardíaca	Inteiro	Preditora
	Temperatura	Ponto Flutuante	Preditora
	Peso	Ponto Flutuante	Preditora
	IMC (Índice Massa Corporal)	Ponto Flutuante	Preditora
	Altura	Inteiro	Preditora
Evolução	Tabagismo	Catégorico	Preditora
	Etilismo	Catégorico	Preditora
	Sulfonilureia	Catégorico	Preditora
	Glinida	Catégorico	Preditora
	Metformina	Catégorico	Preditora
	Pioglitazona	Catégorico	Preditora
	Inibidor de dpp4	Catégorico	Preditora
	Inibidor de SGLT2	Catégorico	Preditora
	Inibidor de alfa-glicosidase	Catégorico	Preditora
Insulina	Catégorico	Preditora	
Acolhimento / Evolução	Glicose	Catégorico	Desfecho

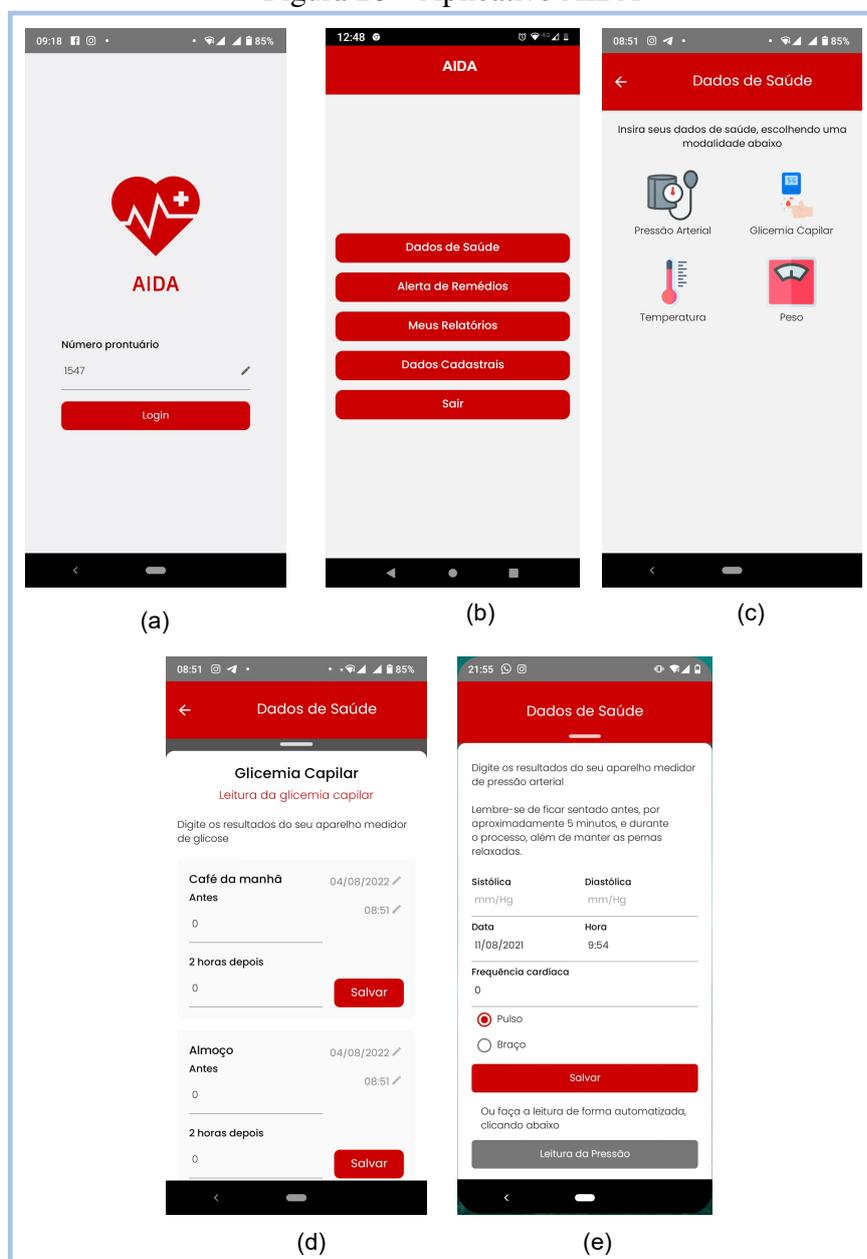
Fonte: Elaborado pelo autor

No total foram utilizadas 19 variáveis preditoras e uma variável de desfecho. Na tabela 13, as variáveis foram agrupadas conforme sua origem, sendo variáveis que se originaram nos dados de acolhimento ao paciente e variáveis que foram recuperadas junto ao campo evolução. Além, destas, a variável alvo (desfecho) foi coletada dos dados de acolhimento e, posteriormente, foi conferida e atualizada através dos dados revisados do campo evolução. Como a variável de interesse possui apenas dois valores possíveis (0, que representa não hipoglicemia e 1 que representa hipoglicemia ou crise hipoglicêmica), caracteriza-se como uma tarefa de classificação binária a ser processada pelos algoritmos de aprendizado de máquina.

5.2.2.4 Dados Coletados no Contexto do Paciente

Como detalhado na Seção 4.3, a Arquitetura AIDA prevê a coleta de dados no contexto do paciente, para posterior inclusão destes dados no RES. Esses dados coletados serviriam para atualizar o modelo de classificação e de predição da arquitetura. Desta forma, foi desenvolvido um aplicativo denominado AIDA app. Este aplicativo tem como objetivo realizar a coleta dos dados dos pacientes em seu contexto. Assim, o aplicativo foi instalado no *smartphone* dos pacientes e estes informavam os valores dos sinais vitais conforme orientações de especialistas.

Figura 26 – Aplicativo AIDA



Fonte: Elaborado pelo autor

Na Figura 26 são apresentadas telas do aplicativo AIDA em operação. A primeira tela

(a) de acesso ao aplicativo insere o número do prontuário do paciente (que se conectará ao RES do paciente); na tela seguinte (b), o menu principal com as opções de inserção de dados de saúde, registro de alerta de medicamentos e outras opções. Na tela (c), podemos ver as opções de entrada de dados, que correspondem aos mesmos dados coletados do paciente durante o atendimento no centro de saúde. Nas telas (d) e (e), temos dois exemplos de entrada de dados correspondentes aos dados de glicemia capilar e pressão arterial, respectivamente, com as opções de entrada manual de dados e leitura automática por meio de sensores.

Os dados que foram coletados via AIDA app, foram armazenados de maneira diferente dos dados coletados no acolhimento ao centro clínico. Na Tabela 14 apresentamos um exemplo de armazenamento de dados coletados via aplicativo. Neste caso, os dados coletados se referem a glicose capilar, realizada seis vezes ao dia.

Tabela 14 – Dados de glicose capilar coletados via AIDA app

id paciente ^a	Data Hora ^b	Sexo ^c	Idade ^d	Tipo ^e	Antes ^f	Depois ^g
...
7	19/01/2022 09:30	M	29.44	c	195	292
7	19/01/2022 12:45	M	29.44	a	234	240
7	19/01/2022 22:00	M	29.44	j	320	281
7	20/01/2022 08:30	M	29.44	c	154	252
7	20/01/2022 13:00	M	29.44	a	150	178
7	20/01/2022 22:58	M	29.44	j	317	263
...

^aid do paciente (serve de conexão ao prontuário); ^bData e hora da refeição; ^cSexo do paciente; ^dIdade do Paciente; ^eTipo de refeição (c - café da manhã, a - almoço, j - jantar); ^fGlicose capilar medida imediatamente antes da refeição; ^gGlicose capilar medida duas horas após a refeição

Fonte: Elaborado pelo autor

Na Tabela 15, apresentamos os dados de peso corporal e tipo de balança que foi utilizada para coleta, idade, sexo, além de data e hora da coleta do dado e identificação do paciente.

Tabela 15 – Dados de peso coletados via AIDA app

id paciente ^a	Data ^b	Hora ^c	Sexo ^d	Idade ^e	Tipo ^f	Peso ^g
...
7	19/01/2022	22:45	M	29.44	a	95.0
7	20/01/2022	22:56	M	29.44	a	96.0
7	22/01/2022	00:48	M	29.44	a	96.0
7	23/01/2022	22:19	M	29.44	a	96.3
7	25/01/2022	23:50	M	29.44	d	96.1
7	26/01/2022	18:30	M	29.44	d	96.0
...

^aid do paciente (serve de conexão ao prontuário); ^bData; ^cHora em que o paciente foi pesado; ^dSexo do Paciente; ^eIdade do Paciente; ^fTipo de balança (a - analógica, d - digital); ^gPeso medido

Fonte: Elaborado pelo autor

Na Tabela 16, apresentamos os dados temperatura medida do paciente, normalmente coletada uma vez ao dia, sem restrição com relação ao horário da coleta.

Tabela 16 – Dados de temperatura coletados via AIDA app

id paciente ^a	Data ^b	Hora ^c	Sexo ^d	Idade ^e	Tipo ^f	Temperatura ^g
...
7	19/01/2022	22:45	M	29.44	d	35.9
7	20/01/2022	22:56	M	29.44	d	36.0
7	22/01/2022	00:48	M	29.44	d	36.6
7	23/01/2022	22:19	M	29.44	d	36.3
7	25/01/2022	23:50	M	29.44	d	35.8
7	26/01/2022	18:30	M	29.44	d	35.7
...

^aid do paciente (serve de conexão ao prontuário); ^bData; ^cHora em que foi medida a temperatura; ^dSexo do Paciente; ^eIdade do Paciente; ^fTipo de balança (a - analógica, d - digital); ^gTemperatura medida

Fonte: Elaborado pelo autor

Por fim, na Tabela 17 aparecem os dados de pressão arterial sistólica e diastólica, além do batimento cardíaco. Foi orientado aos pacientes que realizassem essas medições no mesmo horário do dia, se possível.

Tabela 17 – Leituras de pressão arterial e batimentos cardíacos

id paciente ^a	Data ^b	Hora ^c	Sexo ^d	Idade ^e	PAS ^f	PAD ^g	BC ^h
...
7	19/01/2022	22:45	M	29.44	102	74	84
7	20/01/2022	22:56	M	29.44	108	68	84
7	22/01/2022	00:48	M	29.44	105	69	86
7	23/01/2022	22:19	M	29.44	111	66	74
7	24/01/2022	00:16	M	29.44	096	66	75
7	25/01/2022	23:45	M	29.44	099	65	77
...

^aid do paciente (serve de conexão ao prontuário); ^bData; ^cHora em que foi medidos os valores; ^dSexo do Paciente; ^eIdade do Paciente; ^fPressão arterial sistólica; ^gPressão arterial diastólica; ^hBatimento cardíaco

Fonte: Elaborado pelo autor

Os sinais vitais coletados via aplicativo correspondem aos mesmos sinais vitais coletados no acolhimento ao paciente. A ideia de coletar os mesmos dados é para manter o padrão de funcionamento da arquitetura, no que se refere aos algoritmos de aprendizado de máquina e, com isso, poder realizar a predição.

5.2.3 Pré-processamento de dados

O pré-processamento de dados é uma das fases mais significativas no processo de mineração de dados. Ele prepara e transforma o conjunto de dados inicial, geralmente incompletos e in-

consistentes, para evitar produzir resultados enganosos (FARUQUE et al., 2020). Assim, alguns métodos de pré-processamento de dados podem ser aplicados a dados brutos antes de executar uma análise. O conjunto de dados do estudo é baseado na entrada de dados do paciente (dados do RES + dados coletados via aplicativo), com vários fatores complicadores, incluindo valores ausentes, entradas erradas, recursos calculados e amostragem irregular. Portanto, etapas de pré-processamento são necessárias para limpar os dados e torná-los compatíveis com o modelo de aprendizado de máquina proposto. As etapas de pré-processamento consideradas para este estudo incluem manipulação de valores ausentes, normalização e balanceamento dos dados.

5.2.3.1 Tratamento de dados ausentes

Em geral, medições de saúde armazenadas em banco de dados são amostradas irregularmente e apresentadas com pontos de dados ausentes (FARUQUI et al., 2019). Conjuntos de dados médicos reais geralmente consistem em dados ausentes com diferentes padrões que diminuem o desempenho dos classificadores usados em sistemas inteligentes de saúde e diagnóstico de doenças (NADIMI-SHAHRAKI et al., 2021). É necessário usar métodos de imputação de dados ausentes para preencher estes valores para chegar a resultados experimentais validados, garantindo resultados e previsões imparciais (CENITTA; ARJUNAN; V, 2022). Para lidar com os valores ausentes foram adotadas metodologias diferentes de preenchimento, dependendo do tipo de sinal vital faltante e do motivo pelo qual o dado estava faltante. Na Tabela 18 podemos ver os dados de sinais vitais que possuem dados faltantes e sua porcentagem sobre os totais de dados de cada sinal vital.

Tabela 18 – Tabela com quantidade e porcentagem de dados faltantes

Sinal Vital	Quantidade de Nulos	Porcentagem de Nulos
Pressão Diastólica	47	2,3%
Pressão Sistólica	44	2,1%
Frequência Cardíaca	70	3,4%
Temperatura	80	3,9%
Peso	147	4,3%
IMC ^a	184	4,3%
Altura	66	3,2%

^aÍndice de Massa Corporal, calculado com o uso do peso e da altura.

Fonte: Elaborado pelo autor

Diferentes sinais vitais tiveram diferentes tratamentos para dados faltantes. No caso da altura, foi verificado que existiam inconsistências de informações (mesmo paciente com alturas diferentes) e falta de dados (campo altura sem preenchimento). Nesta situação em específico, foi definido junto a especialistas a substituição dos valores de alturas diferentes ⁸, pelo valor de altura com maior frequência nos dados do paciente. Essa mesma abordagem foi utilizada

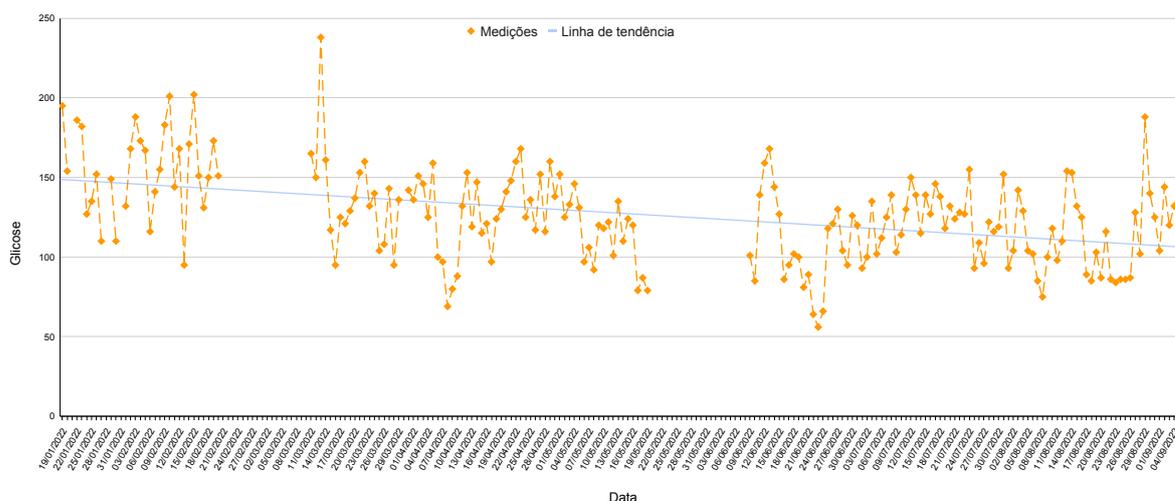
⁸Essa diferença acontece, muitas vezes, porque na hora da coleta dos dados, no acolhimento ao paciente, este pode ser medido de forma diferente, sem o calçado ou com o calçado, por ex.

quando havia dado faltante no campo altura, sendo colocado no campo a altura que aparecia com mais frequência nos dados. Com a abordagem utilizada nas informações de altura, a porcentagem de dados faltantes de IMC reduziu, pois a altura é um dos dados utilizados para o cálculo do IMC.

Para o preenchimento dos demais dados faltantes foi utilizada uma abordagem baseada em imputação múltipla (PEDERSEN et al., 2017; CUMMINGS, 2013; STERNE et al., 2009). O algoritmo *Multiple Imputation by Chained Equation (Mice)* é baseado na *Fully Conditional Specification*, que inicialmente imputa os dados ausentes em cada variável usando a média/modo e, em seguida, imputa cada variável incompleta por um modelo separado, que explora os valores previamente imputados das outras variáveis (BUUREN; GROOTHUIS-OUDSHOORN, 2011; BUUREN, 2018). Na sua configuração padrão, para cada variável com valores ausentes, *mice* usa a parte observada para ajustar uma correspondência de média preditiva (para variáveis numéricas), ou uma regressão logística (para variáveis binárias) ou uma regressão polinômica (para variáveis categóricas) e, em seguida, prevê a parte ausente usando o modelo ajustado (BUUREN, 2018).

Os dados coletados dos paciente através do aplicativo AIDA, também apresentam dados faltantes. O gráfico da Figura 27 apresenta a coleta de dados de glicose capilar do paciente no período de 19/01/2022 a 05/09/2022 e a linha de tendência dos valores. Duas lacunas de falha de coleta de dados, especificamente no intervalo de 21/02/2022 e 10/03/2022 e no intervalo 20/05/2022 e 08/06/2022, podem ser vistas. Para além das lacunas de dados faltantes já citadas, houveram algumas faltas pontuais, distribuídas ao longo do período de coleta. Esses valores ausentes representam cerca de 18% da quantidade total de dados.

Figura 27 – Gráfico com medições de glicose capilar e linha de tendência



Fonte: Elaborado pelo autor

Como o objetivo de tratar os dados faltantes do resultado da coleta via app AIDA, foram utilizadas três abordagens diferentes. A primeira abordagem, que pode ser vista na Figura 28(A),

consiste em repetir os dados coletados imediatamente posteriores à lacuna faltante (MIDRONI et al., 2018), seguindo a indicação da linha de tendência da Figura 27 (que aponta o decréscimo nos valores de leitura ao longo da série histórica). A segunda e terceira abordagem, consistiram em preencher as lacunas de dados faltantes com a média de dados dos quatro dias seguintes, conforme pode ser visto na Figura 28(B), e sete dias seguintes, conforme a Figura 28(C) (MIDRONI et al., 2018; JEON et al., 2020; JAVIDI et al., 2022). Para a imputação dos dados faltantes pontuais, foi utilizada a técnica de imputação múltipla (*mice*) (Duy Le; BEURAN; TAN, 2018).

5.2.3.2 Normalização de Dados

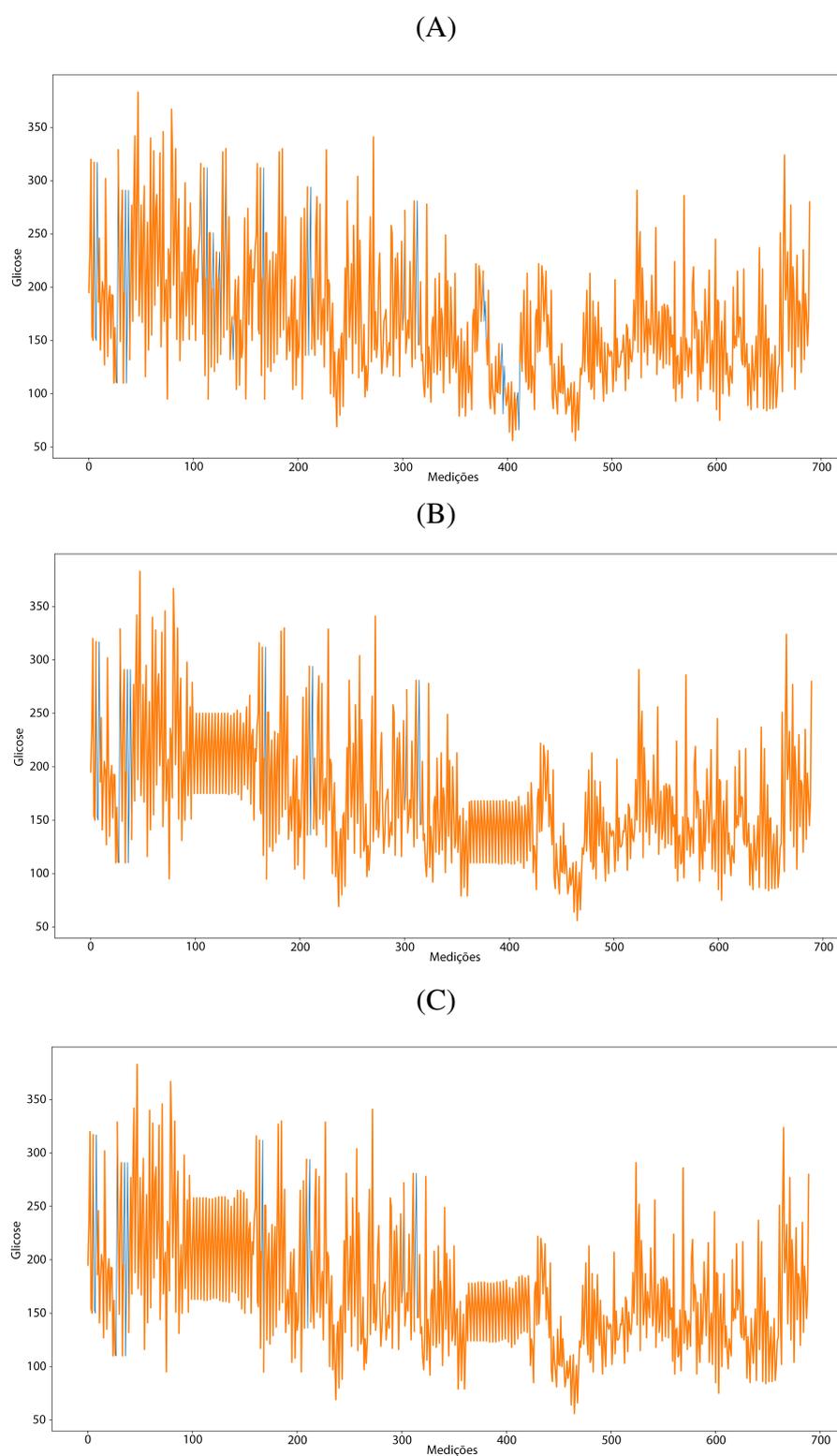
Conforme mostrado na Tabela 11, os valores numéricos estão em uma escala distinta, por exemplo, idade variando de 20 a 70 e altura variando de 150 a 190. Existem variáveis numéricas com diferenças de escala muito grandes (idade, peso, altura pressão arterial, etc) e variáveis categóricas (como sexo). É fundamental unificar seus formatos, para evitar vieses decorrentes de valores muito ampliados (ACHARYA et al., 2017). Para variáveis categóricas, utilizamos a codificação *one-hot-encoding*, utilizando números inteiros simples para lidar com dados textuais, como na coluna sexo. A normalização dos dados, com o objetivo de alterar os valores do conjunto de dados para uma escala comum, sem distorcer diferenças nos intervalos de valores, foi realizada utilizando *standard scaler*. Da mesma forma, conforme demonstrado na Figura 27, os valores de glicemia medidos do paciente possuem uma variação muito grande de amplitude. Assim, para não prejudicar o desempenho do algoritmo de predição foi utilizado como normalização de dados *min-max scaler* (RAKTHANMANON et al., 2013), utilizada em série de dados temporais de saúde (FARUQUI et al., 2019; JAVIDI et al., 2022).

5.2.3.3 Balanceamento dos Dados

Em um cenário ideal de análise, é esperado que as classes presentes em um conjunto de dados estejam similarmente representadas, mas ao trabalharmos com conjuntos de dados e aplicações reais, é mais provável que os dados disponíveis para análise estejam desbalanceados. Diversos estudos publicados na literatura mostram que conjuntos de dados desbalanceados apresentam diversos desafios para a tarefa de classificação (MAHANI; Riad Baba Ali, 2020; ROSHAN; ASADI, 2020; Khalilpour Darzi; NIAKI; KHEDMATI, 2019). Conforme mostrado na seção 5.2.2.3, o *dataset* utilizado neste trabalho apresenta um desbalanceamento perceptível, com a classe minoritária nos casos de hipoglicemia.

A Figura 29 apresenta as abordagens utilizadas para compensar o desbalanceamento. Pode ser visto na Figura 29(A) a distribuição desigual entre a variável alvo (em verde) e não alvo (em vermelho), apresentando 250 e 1821 registros, respectivamente. Na Figura 29(B) é apresentado o resultado do *dataset* com balanceamento utilizando a técnica *Synthetic Minority Oversam-*

Figura 28 – Gráficos das técnicas utilizadas para imputação de dados.

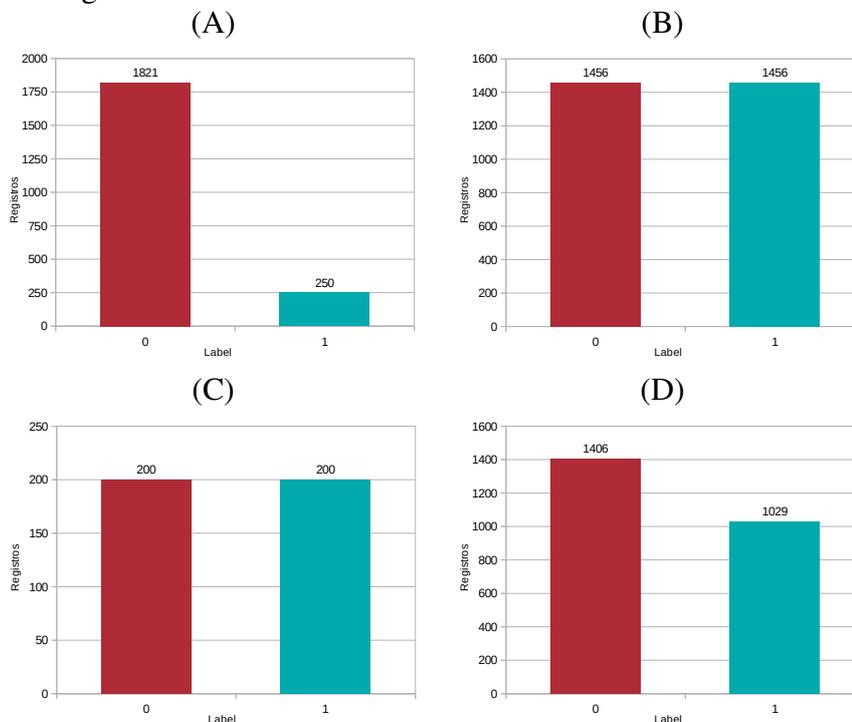


Fonte: Elaborado pelo autor.

pling Technique (SMOTE) - random oversampling, que consiste na replicação de amostras da classe minoritária e, no caso da Figura 29(C) *random undersampling*, que consiste na remoção

aleatória de amostras da classe majoritária (CHAWLA et al., 2002; RAMEZANKHANI et al., 2016; SHUJA; MITTAL; ZAMAN, 2020). Por fim, na Figura 29(D), pode ser visto a técnica *smoteen - combine-sampling* para equilibrar as duas classes de dados (KUMAR et al., 2022).

Figura 29 – Técnicas de Balanceamento de dados utilizadas



Fonte: Elaborado pelo autor.

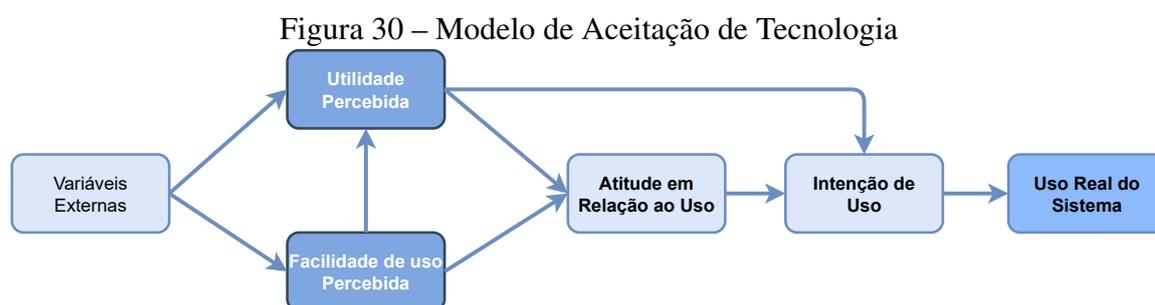
5.3 Métodos de Avaliação

Esta seção discute a avaliação da Arquitetura AIDA, abordando primeiramente a avaliação da tecnologia de coleta de dados AIDA, e sua aceitação por parte dos profissionais de saúde ligados à atenção primária e por parte dos usuários. Em seguida, é realizada a avaliação da classificação e da predição dos algoritmos de aprendizado de máquina aplicados aos dados de históricos dos pacientes e aos dados coletados pelo aplicativo, respectivamente.

5.3.1 Aceitação da Tecnologia

Para implementação e utilização bem sucedida da tecnologia, a vontade e aceitação por parte dos usuários é um fator crucial (AGGELIDIS; CHATZOGLOU, 2009). Durante as últimas décadas, pesquisadores desenvolveram vários modelos para entender os atributos de aceitação da tecnologia entre os usuários. Esses modelos foram verificados várias vezes para determinar sua eficácia para muitas aplicações baseadas em tecnologia da informação, no entanto o modelo de aceitação de tecnologia (TAM) até agora representa a base mais sólida e estabelecida de

aceitação de tecnologia (VENKATESH et al., 2003; CARTER; BÉLANGER, 2005; JAUK et al., 2021). Originário dos campos da sociologia e da psicologia, o TAM é o modelo mais utilizado em diversos estudos de pesquisa. O principal objetivo é prever a adoção de novas tecnologias entre os usuários e destacar os problemas de design do sistema de informação antes que seu uso se torne predominante entre as pessoas. Conforme pode ser visto na Figura 30, o TAM consiste em dois construtos principais: facilidade de utilidade percebida e facilidade de uso percebida que são usados em vários contextos tecnológicos (VENKATESH et al., 2003; YI et al., 2006; AHMAD et al., 2020).



Fonte: Baseado em (DAVIS, 1989)

A utilidade percebida de sistema de informação, pode ser definida como “a extensão em que os indivíduos acreditam que o uso da nova tecnologia aumentará o desempenho de suas tarefas”. Existem muitos estudos na área de tecnologia, que mostram os efeitos da utilidade percebida na intenção de uso de uma determinada forma de tecnologia na área de saúde (KIJSANAYOTIN; PANNARUNOTHAI; SPEEDIE, 2009; RHO et al., 2015; ZHANG et al., 2019). Por sua vez, a facilidade de uso percebida é definida como “o grau em que um indivíduo acredita que o uso de uma determinada tecnologia será livre de esforço mental” (DAVIS, 1989). Muitos estudos no campo da aceitação da tecnologia mostram que a facilidade de uso percebida tem influências diretas e indiretas, mediadas pela utilidade percebida na intenção de usar a tecnologia (SAMPA et al., 2020; BHATIA, 2021; SERRANO et al., 2021; AHMAD et al., 2020).

5.3.1.1 Participação e Avaliação

A amostragem não probabilística de conveniência foi utilizada para selecionar os profissionais de saúde, sendo os participantes escolhidos com base na disponibilidade e não em critérios estatísticos. Também usamos a estratégia bola de neve para selecionar os profissionais de saúde para este estudo, na qual pedimos a cada participante que recomendasse outros candidatos conhecidos. Uma abordagem em dois passos descrita e utilizada em (ANDERSON; GERBING, 1988; NAGLIS; BHATIASEVI, 2019) foi aplicada na fase de avaliação. Na primeira etapa, foram realizados testes para comprovar a confiabilidade da escala por meio da análise do modelo de mensuração. Em seguida, em uma segunda etapa, foi realizada uma análise qualitativa dos resultados obtidos com a aplicação do instrumento, avaliando os resultados das respostas e os

comentários dos respondentes.

5.3.1.2 Instrumento de Avaliação e Intervenção

Como pode ser observado na Tabela 19, uma série de afirmações foram elaboradas e submetidas aos profissionais de saúde para determinar a aceitação e possível adoção da tecnologia para controlar o curso da doença DM2. A população-alvo definida para esta pesquisa foram os profissionais de saúde diretamente envolvidos no acompanhamento e tratamento de pacientes com DM2, a exemplo daqueles que atuam na atenção básica, hospitais e médicos da área. A partir dessas afirmações, foi disponibilizado um instrumento online. O instrumento inicial foi discutido com especialistas e foram feitas as modificações necessárias para melhorar o conteúdo e a clareza. Posteriormente, foi desenvolvido na plataforma Google Forms e apresentado diretamente aos envolvidos por meio de mensagens de e-mail e aplicativos de mensagens. Junto com o formulário de pesquisa, um vídeo apresentando as funcionalidades do aplicativo e um link para download do protótipo foram disponibilizados para os interessados em testar as funcionalidades disponíveis. A pesquisa baseada na internet foi adotada neste estudo devido à flexibilidade demográfica e segurança no alcance do público-alvo, levando em consideração os efeitos da pandemia de Covid-19.

O instrumento era composto por duas partes. A primeira contemplou as características demográficas essenciais dos participantes, incluindo informações sobre sexo, idade, local de trabalho e profissão/escolaridade. A segunda parte continha afirmações para avaliar a aceitação do uso do aplicativo com base nos conceitos do TAM. Assim, foi utilizada a escala Likert de 5 pontos, que variou entre discordo totalmente (1) a concordo totalmente (5). O instrumento foi disponibilizado através da ferramenta *Google Forms*⁹, juntamente com um vídeo explicativo das funcionalidades da tecnologia e do app¹⁰. A coleta de dados teve início em julho de 2021 e foi concluído em agosto de 2021.

5.3.1.3 Análise dos dados

A validade de conteúdo do instrumento de avaliação, foi realizada quantitativamente por meio do Índice de Validade de Conteúdo por Item (I-CVI). Esse índice indica a proporção de concordância sobre a relevância de cada item, considerando como valor mínimo justo um resultado maior ou igual a 0,80. Para avaliar a abrangência do roteiro como um todo, foi verificado o S-CVI/Ave (Índice de Validade de Conteúdo por Escala/Média). Aqui, valores maiores ou iguais a 0,80 são considerados representativos de relevância (POLIT; BECK, 2006; da Silva Lima Roque et al., 2021).

Além disso, para estimar a confiabilidade dos itens de mensuração, o alfa de Cronbach¹¹ foi

⁹<https://forms.gle/VQpoKzoY7WqMz1mM6>

¹⁰<https://youtu.be/DbOkwEbPUD8>

¹¹Coeficiente calculado para medir a consistência interna/confiabilidade de um instrumento de avaliação.

Tabela 19 – Itens dos Construtos

Itens dos Construtos**Utilidade percebida**

- 1 O aplicativo AIDA auxilia no monitoramento do paciente com DM2.
- 2 O aplicativo AIDA possui as funções necessárias para coleta dos sinais vitais e sintomas da DM2.
- 3 O aplicativo AIDA é apropriado para atender aos objetivos da equipe de atenção primária.
- 4 O aplicativo AIDA é apropriado para atender aos objetivos da equipe hospitalar.
- 5 O aplicativo AIDA auxilia o paciente no monitoramento do DM2.

Facilidade de uso percebida

- 6 O AIDA é fácil de utilizar.
- 7 A interface do AIDA facilita a realização do acompanhamento.
- 8 A interface do AIDA facilita a realização do autocuidado.

Intenção de uso

- 9 Usar o app AIDA é uma boa ideia.
- 10 Usar o app AIDA é agradável.

Fonte: Elaborado pelo autor

calculado sobre os construtos. Valores de alfa de Cronbach acima de 0,7 indicam a confiabilidade do instrumento e, como forma complementar, foi calculada a correlação de Pearson entre os itens dos construtos (NATARAJAN; BALASUBRAMANIAN; KASILINGAM, 2018). A matriz de correlação mede o grau de relação linear entre cada par de itens, podendo ficar entre -1 e 1 (BAABDULLAH, 2018). Por fim, uma análise qualitativa foi realizada, para avaliar os resultados das respostas do questionário e os comentários dos respondentes.

5.3.1.4 Entrevista de avaliação com os pacientes

Com o propósito de complementar os dados obtidos na pesquisa sobre a aceitação da tecnologia entre os profissionais envolvidos no acompanhamento de pacientes com Diabetes Mellitus tipo 2 (DM2), realizou-se uma entrevista com os próprios pacientes que utilizaram o aplicativo AIDA. A abordagem da entrevista foi do tipo centralizada, na qual o entrevistador permite que o entrevistado expresse livremente sua experiência pessoal no uso do aplicativo, dentro de tópicos predefinidos, conforme preconizado por (LAKATOS; MARCONI, 2010; BARROS; LEHFELD, 2007). Para essa finalidade, foram formuladas uma série de perguntas com base nas

afirmações presentes no instrumento de avaliação de usabilidade conhecido como SUS (System Usability Scale), conforme descrito por (MARTINS et al., 2015). As perguntas com base no SUS estão detalhadas na Tabela 20.

Tabela 20 – Questionário baseado nas afirmações do SUS

Afirmação Original SUS	Pergunta Adaptada
Acho que gostaria de utilizar este produto com frequência.	Você usaria o aplicativo se ele estivesse disponível (ao invés da folha de anotação)?
Considerarei o produto mais complexo do que necessário.	Achou que o aplicativo poderia ser mais fácil de utilizar?
Achei o produto fácil de utilizar.	Achei o aplicativo fácil de utilizar?
Acho que necessitaria de ajuda de um técnico para conseguir utilizar este produto.	Precisou de auxílio para utilizar o aplicativo (cuidador ou técnico)?
Considerarei que as várias funcionalidades deste produto estavam bem integradas.	Achou as funcionalidades do produto integradas? Fluiu o uso
Achei que este produto tinha muitas inconsistências.	Encontrou muitos problemas (inconsistências - instabilidade) no uso do aplicativo?
Suponho que a maioria das pessoas aprenderia a utilizar rapidamente este produto.	Acha que outros pacientes aprenderiam a utilizar o aplicativo rapidamente e com facilidade?
Considerarei o produto muito complicado de utilizar.	Achou o aplicativo muito complicado de ser utilizado?
Senti-me muito confiante a utilizar este produto.	Qual a sensação que teve ao utilizar o aplicativo (confiança, segurança)?
Tive que aprender muito antes de conseguir lidar com este produto.	Teve que estudar/aprender muito para utilizar o aplicativo?

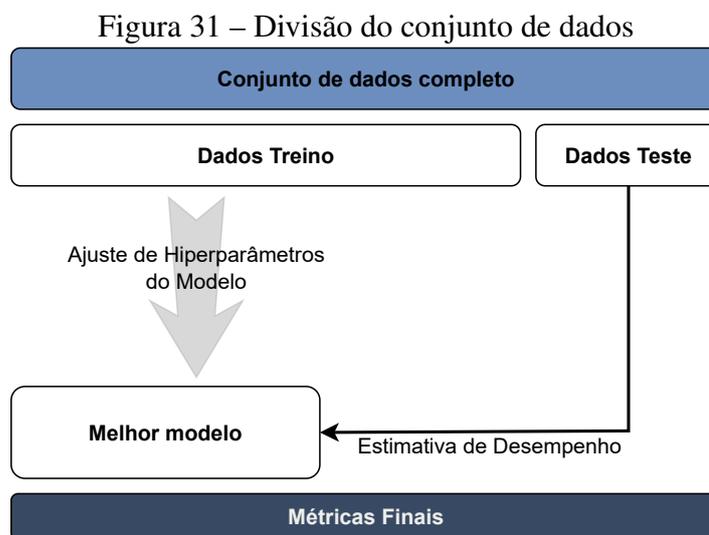
Fonte: Baseado em (MARTINS et al., 2015)

Adicionalmente, é importante salientar que as questões orientadoras das entrevistas com os pacientes, conforme apresentadas na Tabela 20, foram direcionadas para compreender a relevância do uso do aplicativo no processo de gestão da doença, na coleta de informações relevantes, bem como na percepção global em relação ao aplicativo. Ressaltando-se que as entrevistas foram gravadas em meio digital.

5.3.2 Treinamento e teste

Ao trabalhar com aprendizado de máquina, usamos métricas específicas para avaliar qual algoritmo melhor se adapta à tarefa e ao conjunto de dados. Sempre que possível, para testar o modelo obtido, utiliza-se um conjunto de dados de validação externa, coletado independentemente do conjunto de dados de treinamento no qual o modelo foi construído (REDDY; AGGARWAL, 2015). No entanto, na maioria dos casos práticos, os dados são escassos e difíceis de serem coletados. Para resolver este problema, uma estratégia comum é dividir o conjunto de dados em conjuntos de treinamento e teste, antes de ajustar um modelo de aprendizagem de máquina e prosseguir com a avaliação final do melhor desempenho do modelo. A razão para fazer

isso é que, se usarmos todo o conjunto de dados para ajuste, o modelo sofrerá um *sobreajuste* dos dados e poderá levar a previsões ruins em cenários futuros. Portanto, manter uma parte do conjunto de dados e testar o desempenho do modelo antes de implantá-lo, pode proteger contra problemas inesperados que podem surgir devido ao *overfitting* (JOSEPH, 2022). A Figura 31 ilustra a divisão do conjunto de dados em dados de treinamento e teste.



Fonte: Elaborado pelo autor

Não há consenso sobre a proporção ideal de divisão de dados para treinamento e teste. As divisões mais usadas são 60:40, 70:30 ou 80:20, dependendo do tamanho do conjunto de dados inicial. Para grandes conjuntos de dados, o conjunto de teste pode ser apenas 10% ou menos (RASCHKA et al., 2022; JOSEPH; VAKAYIL, 2022). Neste estudo utilizamos a divisão de proporção 80:20. O Capítulo 6 apresenta mais detalhes e os resultados de cada análise realizada. Para conseguir essa divisão, usamos a função `train_test_split`¹² do Scikit-learn, configurada para embaralhar o conjunto de dados antes de realizar a divisão estratificada e permitir que a divisão seja reproduzida em vários testes. Estratificação significa que o método `train_test_split` retorna subconjuntos de treinamento e teste que têm a mesma proporção de rótulos de classe que o conjunto de dados de entrada.

5.3.3 Seleção do Modelo

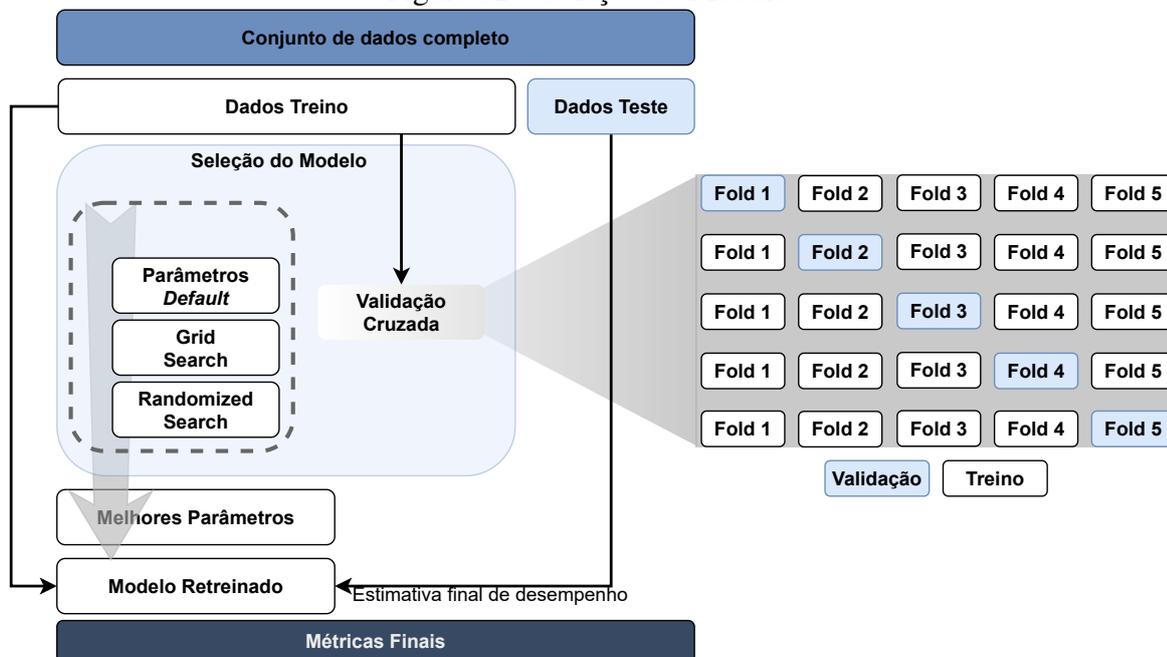
Cada algoritmo de classificação tem vieses inerentes e nenhum modelo de classificação pode ser considerado superior, se não fizermos suposições sobre a tarefa. Na prática é essencial comparar diferentes algoritmos de aprendizado para treinar e selecionar o modelo de melhor desempenho (RASCHKA et al., 2022). Um modelo de aprendizado de máquina é composto de parâmetros e hiperparâmetros que afetam a velocidade e a precisão do processo de aprendizado. Um parâmetro é uma variável cujo valor é estimado a partir do conjunto de dados e

¹²https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html

aprendido durante o treinamento dos conjuntos de dados históricos. Em contraste, um hiperparâmetro é uma variável de configuração externa ao modelo que é definida manualmente antes do treinamento do modelo, sendo que seu valor não pode ser avaliado a partir dos conjuntos de dados.

Nesta seção, apresentaremos a abordagem utilizada para encontrar o melhor modelo de ML para a tarefa e dados disponíveis. Um modelo é constituído pela configuração dos diferentes hiperparâmetros, com cada algoritmo de ML precisa de um número diferente de hiperparâmetros. O desafio é encontrar os melhores valores para os hiperparâmetros. Não é possível saber o melhor conjunto de valores do hiperparâmetro, mas podemos usar regras práticas ou selecionar um valor com tentativa e erro para nosso algoritmo de ML. A Figura 32 ilustra nossa abordagem para encontrar a combinação apropriada de hiperparâmetros para treinar o melhor modelo de cada algoritmo de aprendizado de máquina selecionado.

Figura 32 – Seleção do Modelo



Fonte: Baseado em (SCIKIT-LEARN, 2022a)

Na Seção 5.3.2 apresentamos a necessidade da divisão inicial do conjunto de dados em dados de treinamento (usados para ajustar o modelo de aprendizado de máquina) e dados de teste (usados para avaliar o desempenho do melhor modelo). Usamos as abordagens de otimização de dados de treinamento e hiperparâmetros nesta etapa para ajustar os modelos. Para isso, o treinamento é dividido em múltiplos conjuntos, e o modelo candidato foi treinado e validado usando o procedimento de validação cruzada (JOSEPH; VAKAYIL, 2022). Optamos por usar uma validação cruzada de K -Folds com cinco *fold*s.

Na validação cruzada, o conjunto de dados é dividido aleatoriamente em um número de k -*fold*s (porções de dados) de tamanho igual. Em seguida, cada *fold* é utilizada como um conjunto de teste e as outras são usadas para treinamento. Cada iteração usa um *fold* diferente como teste,

o que produz uma estimativa de erro. Após completar o teste em um *fold* e utilizado os outros como treinamento, um *fold* sucessivo diferente do anterior, é utilizado, e o procedimento é repetido para produzir outra estimativa de erro. O processo continua até que todas os *k-fold* sejam usados uma vez como um conjunto de teste. Assim, temos um número *k* de estimativas de erro que podem ser computadas em uma estimativa de erro médio (a pontuação de validação cruzada) e um erro padrão das estimativas (RASCHKA; MIRJALILI, 2019).

Não podemos esperar que os hiperparâmetros padrão dos diferentes algoritmos de aprendizado fornecidos pelas bibliotecas de software seja ideal para nossa tarefa de problema específica. Portanto, faremos uso frequente de técnicas de otimização de hiperparâmetros que nos ajudam a ajustar o desempenho do nosso modelo (RASCHKA; MIRJALILI, 2019). Para fins de comparação, treinamos um modelo com a configuração padrão dos hiperparâmetros do Scikit-learn para cada um dos algoritmos testados. A biblioteca *Scikit-learn* fornece duas abordagens genéricas para otimização de hiperparâmetros, o *Grid Search* e o *Randomized Search* (SCIKIT-LEARN, 2022b).

O *Grid Search* considera exaustivamente todas as combinações de parâmetros para gerar candidatos a partir de uma grade de valores de parâmetros especificados anteriormente. Essa estratégia de busca possui um custo computacional significativo que deve ser considerado em conjuntos de dados com muitas características, mas atualmente é o método mais utilizado para otimização de parâmetros. Por sua vez, o *Randomized Search* implementa uma pesquisa aleatória sobre os hiperparâmetros, onde cada configuração é amostrada de uma distribuição sobre possíveis valores de hiperparâmetros. Esta técnica tem dois benefícios principais sobre uma busca exaustiva: uma combinação pode ser escolhida independente do número de hiperparâmetros e valores possíveis, e adicionar hiperparâmetros que não influenciam o desempenho não diminui a eficiência.

Ao final do processo de otimização de hiperparâmetros, temos três modelos de ML que foram ajustados no dataset de treinamento: um com hiperparâmetros padrão, um gerado por *Grid Search* e outro como resultado de *Randomized Search*. Os modelos são treinados novamente com todo o conjunto de treinamento e usamos o conjunto de dados de teste para estimar o desempenho dele nesses dados não vistos. Se o desempenho for satisfatório, o modelo pode ser usado para prever novos dados futuros (RASCHKA; MIRJALILI, 2019).

5.3.4 Parâmetros de desempenho

Avaliar a qualidade é uma parte importante do desenvolvimento do modelo. Para mensurar os resultados apresentados pelo AIDA serão avaliados a precisão do modelo para detecção de crises e a precisão do modelo em relação a predição de crises. Como forma de quantificar a qualidade da arquitetura, separamos as métricas de classificação e regressão em duas subseções.

5.3.4.1 Métricas para quantificar a qualidade da classificação

Em um problema de classificação binária, as instâncias são mapeadas para duas classes, por exemplo, positivas ou negativas. Formalmente, um classificador X mapeará as instâncias para uma das duas classes positivas ou negativas $\{p, n\}$. Assim, dado qualquer classificador binário, existem quatro resultados possíveis (FAWCETT, 2006; NICOLAS, 2017):

Verdadeiro Positivo (VP): Essas são as observações corretamente rotuladas como pertencentes à classe positiva;

Verdadeiro Negativo (VN): Essas são observações incorretamente rotuladas como pertencentes à classe positiva;

Falso Positivo (FP): Essas são observações incorretamente rotuladas como pertencentes à classe positiva;

Falso Negativo (FN): Essas são observações incorretamente rotuladas como pertencentes à classe negativa.

Estes fatores podem ser utilizados para avaliar acurácia, precisão, sensibilidade e *F-score* (que é uma medida que combina precisão e sensibilidade).

Acurácia: É a base da medição da qualidade de qualquer modelo preditivo. A acurácia mede a capacidade do modelo de classificar corretamente cada caso de o paciente estar em crise ou não estar em crise, dado o número total de casos avaliados. Na equação 5.1, ac fornece a medida para a acurácia.

$$ac = \frac{VP + VN}{VP + VN + FP + FN} \quad (5.1)$$

Precisão: A precisão de um modelo é a fração de ocorrências relevantes entre as ocorrências recuperadas. É calculado considerando a razão de verdadeiros positivos pelo total de positivos em um modelo. No caso do modelo AIDA significa dizer a quantidade de crises detectadas de forma correta. Em 5.2 apresentamos a equação para a precisão, dado por p .

$$p = \frac{VP}{VP + FP} \quad (5.2)$$

Sensibilidade: É a fração de ocorrências relevantes que foram recuperadas sobre a quantidade total de ocorrências relevantes. Uma sensibilidade alta significa que a maioria das ocorrências retornadas eram relevantes, é a capacidade que o modelo tem de detectar se a pessoa

está em crise. É medido como a razão de verdadeiros positivos para a soma dos verdadeiros positivos e falsos negativos dado por s , apresentada em 5.3

$$s = \frac{VP}{VP + FN} \quad (5.3)$$

$F_1 - Score$: É uma medida que combina precisão e sensibilidade. É aproximadamente a média dos dois quando estão próximos, senão os seus meios harmônicos. A média harmônica é a razão entre o quadrado da média geométrica dividido pela média aritmética. Na medida F_1 , tanto a precisão quanto a recuperação são igualmente ponderadas, conforme definido em 5.4 :

$$F_1 = 2 * \frac{p * s}{p + s} \quad (5.4)$$

O cálculo da precisão e do fator de recuperação para um classificador com mais de duas classes requer a extração e manipulação da matriz de confusão. Porém, essas medidas de avaliação comumente usadas (recall, precisão e $F-score$) são tendenciosas e não devem ser usadas sem uma compreensão clara dos vieses e a identificação correspondente dos níveis de chance ou de caso base da estatística (POWERS, 2011). Uma medida mais confiável da qualidade de um modelo é conhecida como a área sob a curva (AUC - *Area Under the curve*). Uma das mais usadas é a área sob as características operacionais do receptor (AUROC - *Area under the receiver operating characteristic*) (NICOLAS, 2017). Nas ciências médicas, a análise ROC tornou-se um padrão para avaliação e estabelecimento de padrões, comparando a verdadeira taxa positiva e taxa de falso positivo (REDDY; AGGARWAL, 2015).

A curva ROC é uma ferramenta para visualizar o desempenho de um classificador binário. A curva é gerada calculando os valores da taxa positiva verdadeira (VPR) (equação 5.5) e da taxa positiva falsa (FPR) (equação 5.6) para diferentes limites de classificação em um conjunto de validação. As equações para calcular o VPR e o FPR são apresentadas a seguir:

$$TPV = \frac{VP}{VP + FN} \quad (5.5)$$

$$TPF = \frac{FP}{FP + VN} \quad (5.6)$$

Sendo um gráfico bidimensional no qual a taxa de verdadeiros positivos é plotada no eixo Y

e a taxa de falsos positivos é plotada no eixo X. A área sob a curva ROC calcula a integral sob a própria curva, e pode ser interpretada como a probabilidade mais alta de o classificador binário classificar uma observação positiva escolhida aleatoriamente, do que uma observação negativa escolhida aleatoriamente (NICOLAS, 2017).

No domínio biomédico, a curva ROC tem sido empregada na avaliação do diagnóstico da doença. O ponto de corte varia do valor mais alto possível, onde todos os sujeitos são previstos como negativos (TPV = 0, TPF = 0), até o valor mais baixo possível, onde todos os sujeitos são previstos como positivos (TPV = 1, TPF = 1), e em cada corte possível, o TPV e o TPF são calculados com base na matriz de confusão correspondente (REDDY; AGGARWAL, 2015; NICOLAS, 2017).

Em um modelo ideal, onde TPV = 1 e TPF = 0, a área sob a curva ROC (AUC) será igual a 1. A AUC de um classificador aleatório é 0,5. Quanto maior o valor de uma AUC acima de 0,5, melhor o modelo de previsão. Contudo, quando a AUC é menor que 0,5, não significa que o modelo de previsão seja ruim, mas significa que a suposição feita pelo modelo está incorreta e, portanto, para resolver esse problema, basta trocar a definição de indivíduo positivo e indivíduo negativo (REDDY; AGGARWAL, 2015).

5.3.4.2 Métricas para quantificar a qualidade da regressão

Assim como na classificação, a qualidade da regressão é mensurada pelo erro. Contudo, para mensurar o erro na regressão precisamos mensurar a distância entre o ponto predito e o ponto real, também chamada de função de perda. A função de perda mais básica, é a soma dos erros em cada iteração. O erro será a diferença entre o valor previsto e o valor real. Portanto, a função de perda será dada como $L = \sum(\hat{Y} - Y)$ onde L é a perda, \hat{Y} é o valor predito e Y é o valor real (WANG; BOVIK, 2009).

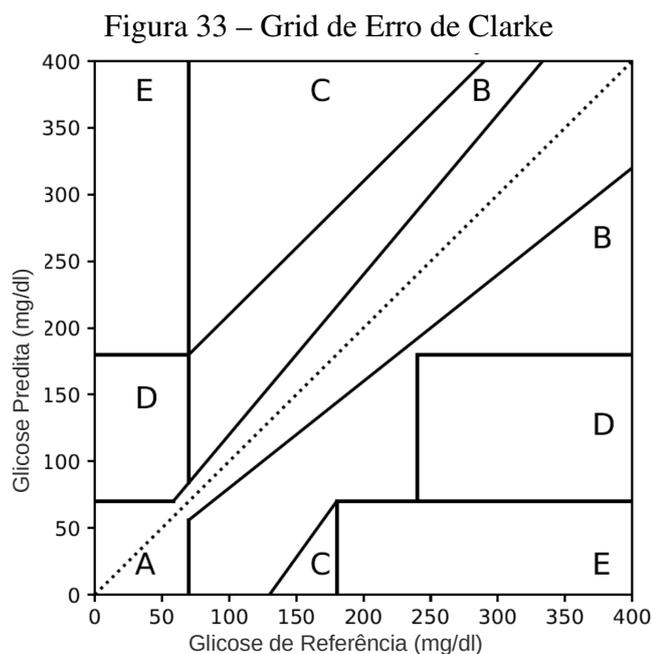
Em trabalhos relacionados à diabetes para predição de glicemia (FARUQUI et al., 2019; KALITA; MIRZA, 2022), biomarcadores ligados ao DM2 (RAMAZI et al., 2019) ou mesmo a predição do surgimento do DM2 (PRIYADARSHINI; BARIK; DUBEY, 2018), a qualidade da regressão é mensurada utilizando funções como RMSE (*Root Mean Squared error*), definido pela Função 5.7, a função MAE (*Mean Absolute Error*), apresentada na Função 5.8 ou ainda a Função MPD (*Mean Poisson Deviance*), que pode ser visualizada na Função 5.9.

$$RMSE(y, \hat{y}) = \sqrt{\frac{\sum_{i=0}^{N-1} (y_i - \hat{y}_i)^2}{N}} \quad (5.7)$$

$$MAE(y, \hat{y}) = \frac{\sum_{i=0}^{N-1} |y_i - \hat{y}_i|}{N} \quad (5.8)$$

$$MPD = 2 \sum y_i \log \left(\frac{\hat{\lambda}}{y_i} \right) - \sum (\hat{\lambda} - y_i) \quad (5.9)$$

O RMSE é o resultado do erro obtido no treinamento do modelo, em geral esse valor representa a distância estimada pelo modelo em relação ao valor obtido. Quanto menor o RMSE, mais próximo do valor real e mais preciso é o modelo treinado. Por sua vez o MAE pode ser definido como a soma absoluta das diferenças entre as variáveis alvo e previstas mede a média da diferença absoluta entre os valores previstos pelo modelo e os valores observados. Por fim, o MDP mede a Média do Desvio de *Poisson*, o objetivo de utilizar esta métrica é de não penalizar caso haja uma tendência de valores preditos acima da média.



Fonte: Baseado em (CLARKE et al., 1987)

As Funções 5.7, 5.8 e 5.9, visualizadas acima, são comumente utilizadas para avaliar o desempenho de modelos de previsão. No entanto, esses critérios não consideram o impacto clínico do erro de previsão e como isso pode afetar a tomada de decisão médica (FARUQUI et al., 2019). Para tanto, iremos considerar outro critério relacionado ao erro quadrático médio, denominado Erro de Grade de Clarke (*Clarke Error Grid*) (CLARKE et al., 1987), para podermos determinar um erro aceitável para a precisão da previsão da glicemia, em comparação com a observação real. O Clark Error Grid é uma das ferramentas mais amplamente utilizadas para avaliar a precisão clínica da estimativa da glicemia (FARUQUI et al., 2019). Conforme pode ser visualizado na figura 33, consiste em um gráfico com cinco zonas principais de atenção (A, B, C, D e E) para interpretação dos níveis de glicose previstos. A Zona A representa aqueles valores dentro de 20% do valor de referência que geralmente levam ao tratamento adequado dos pacientes. A zona B representa aqueles valores que estão fora da zona A, mas não levam ao tratamento inadequado dos pacientes. Os valores de previsão que caem na zona C levam a um tratamento inadequado, mas sem consequências perigosas para o paciente. Os valores de previsão na zona D levam à falha na detecção de hipoglicemia ou hiperglicemia. Finalmente, os valores de predição na zona E levam ao tratamento inadequado da hiperglicemia em vez da

hipoglicemia e vice-versa, dependendo da localização da zona.

5.4 Considerações Parciais

Este capítulo apresentou os materiais e a metodologia utilizada nesta Tese de Doutorado. Foram apresentados uma caracterização da origem dos dados e como ocorreu a organização para sua utilização. Apresentamos a metodologia adotada para o pré-processamento dos dados, englobando métodos utilizados para tratamento de dados faltantes e para a normalização. O app AIDA, que foi utilizado para coleta de dados do paciente em seu contexto foi apresentado, bem como os dados coletados através do aplicativo. A seção sobre métodos de avaliação da arquitetura AIDA aborda a participação e avaliação por parte dos profissionais de saúde envolvidos na atenção primária e por parte dos pacientes DM2. Foi descrito um instrumento de avaliação de aceitação da arquitetura baseado no TAM (*Technology Acceptance Model*) com avaliações qualitativas dos resultados. Com relação aos pacientes, realizamos entrevistas semi-estruturadas para compreender as impressões da utilização da tecnologia e sua implicação no contexto do paciente. Buscamos melhorar a compreensão desses dois grupos sobre aplicação da tecnologia para o acompanhamento do curso da DM2. Por sua vez, a avaliação do desempenho da arquitetura se concentrou no módulos de classificação e de predição. Usando os dados do registro eletrônico dos pacientes para treinar a classificação e os dados coletados no contexto do paciente para treinamento da predição.

6 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados e a discussão dos experimentos realizados para avaliar a arquitetura AIDA. O material e a metodologia utilizada para realizar esses experimentos estão descritos no Capítulo 5. O capítulo está dividido em seis seções. Iniciamos com a avaliação da aceitação da tecnologia para uso com acompanhamento de pacientes diabéticos tipo 2 focada nos profissionais ligados ao atendimento primário em saúde, e após uma avaliação qualitativa dos resultados obtidos com a aplicação do instrumento de avaliação, descritos na Seção 6.1. Após, na Seção 6.2, apresentamos e discutimos a avaliação do modelo de classificação para detectar crises de hipoglicemia e, na Seção 6.3 avaliamos os resultados dos modelos de predição utilizados. Os experimentos descritos nas Seções 6.2 e 6.3 foram realizados no ambiente *Colaboratory* do Google¹. Na seção 6.4 apresentamos a discussão dos resultados obtidos tanto na aceitação da tecnologia quanto nos experimentos de classificação e predição. As limitações do trabalho são apresentados na Seção 6.5 e, por fim, na Seção 6.6 as considerações parciais do capítulo são apresentadas.

6.1 Avaliação de aceitação da Tecnologia

Esta Seção descreve a avaliação da aceitação de parte da arquitetura proposta, correspondente ao aplicativo desenvolvido para coleta de dados dos pacientes DM2. Conforme citado na Seção 5.3.1.1, uma abordagem em duas etapas foi aplicada neste estudo. Na primeira etapa, foram realizados testes para comprovar a confiabilidade da escala por meio da análise do modelo de medição. Em seguida, em uma segunda etapa, foi realizada uma análise qualitativa dos resultados obtidos com a aplicação do instrumento, avaliando os resultados das respostas e os comentários dos respondentes.

6.1.1 Avaliação do instrumento de medida

Para avaliar a confiabilidade do instrumento, alfa de Cronbach foi calculado para garantir um nível adequado de confiabilidade dos construtos. Os resultados apresentados na Tabela 21 indicam que todas as construções atingiram um valor acima do ponto de corte (0,70), variando de 0,75 a 0,88, o que indica que os construtos tem uma robusta confiabilidade interna. O Índice de Validade de Conteúdo por Item (I-CVI) foi calculado para todos os itens, conforme pode ser visto na Tabela 21, atingindo valores, senão acima, próximos a 0,80, considerando-os representativos de relevância. Para complementar o I-CVI, foi verificado o S-CVI (Índice de Validade de Conteúdo por Escala/Média), como forma de avaliar a abrangência do roteiro como um todo. Obtendo um valor de 0.87, confirmando a relevância mostrada de forma individual pelos itens.

Além disso, a correlação de Pearson foi calculada para mostrar o grau de relação linear entre

¹<https://colab.research.google.com/>

Tabela 21 – Índice de validade de conteúdo por item, por escala e alfa de cronbach dos construtos

Construto	Itens	I-CVI	Alfa de Crombach
PU^a	Item 1	0.81	0.88
	Item 2	0.69	
	Item 3	0.81	
	Item 4	0.75	
	Item 5	0.94	
PeoU^b	Item 6	1.00	0.72
	Item 7	1.00	
	Item 8	0.88	
BIU^c	Item 9	0.94	0.85
	Item 10	0.88	
S-CVI/Ave		0.87	

^aUtilidade percebida; ^bFacilidade de uso percebida; ^cIntenção de uso

Fonte: Elaborado pelo autor

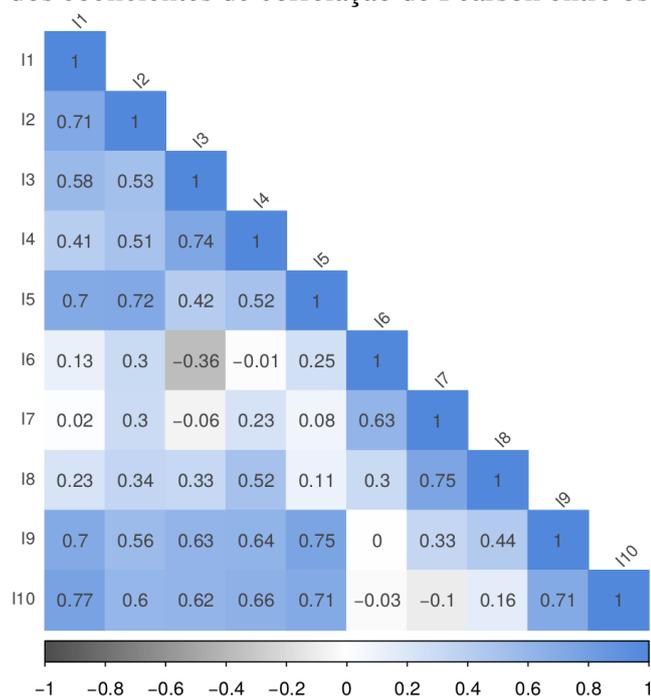
os itens. A Figura 34 mostra que houve correlações positivas significativas entre os itens dentro de seus respectivos construtos. Por exemplo, dentro do construto Utilidade percebida (I1 a I5), as correlações variaram de 0,41 a 0,74, demonstrando uma correlação moderada a forte entre os itens desse mesmo construto. Por sua vez, dentro do construto Facilidade de uso percebida (I6 a I8), as correlações variam de 0,3 a 0,75, demonstrando uma correlação moderada a forte. Por fim, no construto Intenção de uso (E9 e E10), a correlação é de 0,71, demonstrando forte correlação entre esses itens.

6.1.2 Avaliação qualitativa

O instrumento de avaliação, detalhado na Tabela 19, forneceu uma escala *Likert* de 5 pontos (1 a 5), onde um significa discordo totalmente e cinco concordo totalmente. Na Tabela 22 são detalhadas as informações demográficas da população que realizou a avaliação da aplicação da tecnologia. Os participantes consistiam em 17 profissionais da área de saúde, 12 (70.59%) era do sexo feminino e cinco (29.42%) do sexo masculino. A maioria dos respondentes (58.8%) estavam na faixa etária de 30 a 49 anos, a profissão da maioria dos respondentes (58.8%) é de médicos, seguidos pela profissão de enfermagem (29.4%). O local de trabalho da maioria dos respondentes é na atenção primária a saúde (47%), sendo que os demais (53%) se dividem entre hospitais, consultórios médicos e outros (Universidade, ambulatório, etc).

Um mapa de calor dos resultados da pesquisa com os 17 usuários é mostrado na Figura 35. Treze usuários (76,47%) concordam ou concordam totalmente que o aplicativo AIDA auxilia no acompanhamento do paciente com DM2; 12 usuários (70,59%) concordam ou concordam to-

Figura 34 – Análise dos coeficientes de correlação de Pearson entre os itens dos construtos



Fonte: Elaborado pelo autor

talmente que o aplicativo AIDA possui as funções necessárias para a coleta de sinais e sintomas vitais do DM2. Por outro lado, para esses dois itens, três usuários (17,65%), respectivamente, tiveram opinião neutra, não concordando nem discordando e 1 (5,88%) e 2 (11,76%) usuários, respectivamente, discordaram totalmente do auxílio no monitoramento e que o aplicativo possui as funções necessárias para coletar os sinais e sintomas do DM2.

Para os três itens a seguir, que se referem aos objetivos da equipe de atenção básica (item 3), funcionários do hospital (item 4) e a assistência oferecida ao paciente no acompanhamento (item 5), temos 13 usuários (76,47%) concordo fortemente ou concordo com as afirmações dos itens 3 e 4, e 15 usuários (88,24%) com o item 5. Para o item 3, dois usuários (11,76%) discordam ou discordam totalmente com a afirmação, e para os itens 4 e 5, apenas um usuário (5,88%) para cada item. Os demais usuários, dois (11,76%) para o item 3, três (17,65%) para o item 4 e um (5,88%) para o item 5, têm opinião neutra.

Em relação aos itens do construto facilidade de uso percebida, temos 17 usuários (100%) que concordam ou concordam totalmente que o AIDA é fácil de usar e que a interface facilita o acompanhamento. Por sua vez, 15 usuários (76,47%) concordam ou concordam totalmente que a interface AIDA facilita o autocuidado e, para esta última afirmação, apenas dois usuários tiveram uma percepção neutra. Por fim, com relação ao constructo Behavioral intention to use, temos 15 usuários (88,24%) que concordam ou concordam fortemente com a afirmação que usar o AIDA é uma boa idéia e que usar o app AIDA é agradável, como contraponto 1 usuário (5,88%) discorda de que utilizar o AIDA é uma boa idéia e 3 usuários tem uma opinião neutra com relação a estes dois itens.

Tabela 22 – Informação Demográfica

Demografia	Frequencia	Porcentagem
Gênero		
Feminino	12	70.59
masculino	5	29.41
Idade		
20-29	4	23.53
30-39	7	41.17
40-49	3	17.65
50-59	3	17.65
Profissão		
Enfermeira	5	29.42
Clínico Geral	6	35.30
Endocrinologista	3	17.64
Epidemiologista	1	5.88
Outros ^a	2	11.76
Local de trabalho		
Atenção primária	8	47.05
Hospital	3	17.65
Consultório	3	17.65
Outros ^b	3	17.65

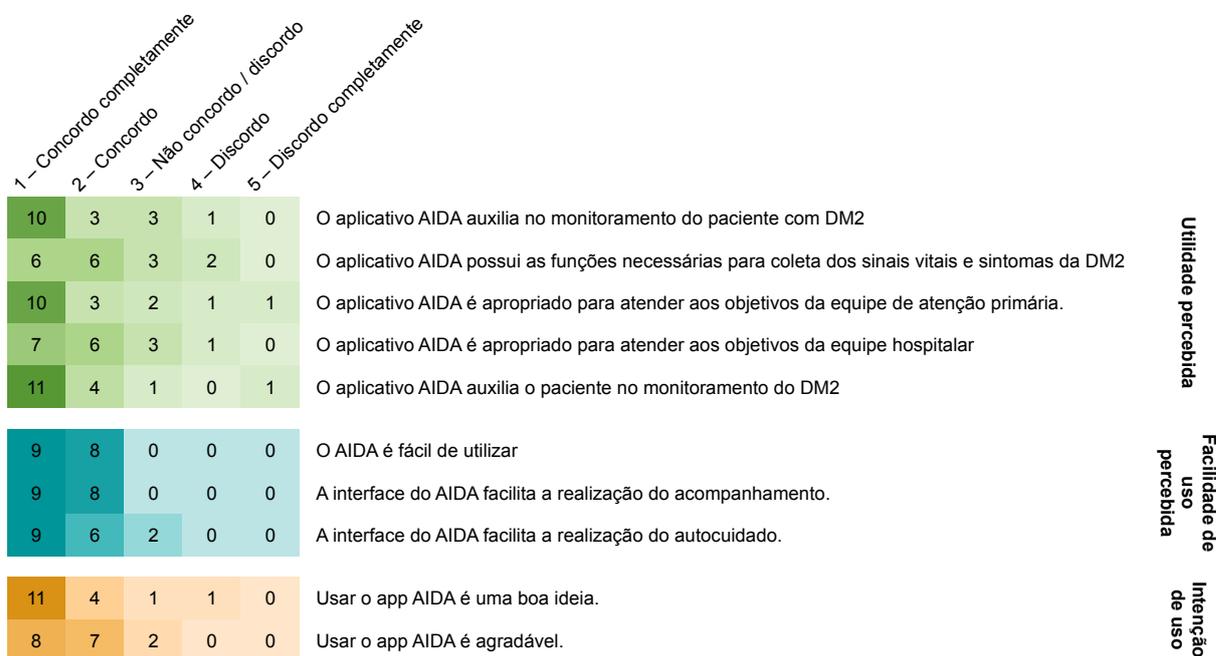
^aEnsino, entre outros; ^bUniversidade, ambulatório

Fonte: Elaborado pelo autor

No geral, os usuários avaliaram bem os três construtos, com maior ênfase na facilidade de uso percebida, com uma avaliação favorável em torno de 96%. Se olharmos para os itens desta construção, percebemos que os profissionais (independentemente do seu local de trabalho) entendem o app AIDA como uma ferramenta fácil de usar. Nesse sentido, dois usuários deixaram comentários no campo de texto livre. O usuário A descreveu o aplicativo como “... uma excelente ferramenta para auxiliar no cuidado e monitoramento de pacientes com DM tipo 2. Considero válida a sua utilização nos serviços de atenção primária à saúde”. O usuário B comentou que “... a preocupação com a facilidade de uso do aplicativo é muito importante. Mas também é preciso levar em conta a carga de trabalho diária que o paciente estaria sujeito a ter para abrir o aplicativo”

Da mesma forma, os constructos *Perceived usefulness* e *Behavior intention to use* também tiveram boas avaliações, com o primeiro alcançando avaliação positiva em torno de 77% e o segundo, em torno de 88%. Na avaliação desses dois construtos, usuários deixaram comentários no campo de texto livre, como: “... já é tempo de inserirmos a tecnologia no cuidado de doenças crônicas” e “Necessário aliar, cada vez mais, tecnologia com os cuidados em saúde. No sentido de ampliar e abranger a população, como no caso dos diabéticos, contribuindo para a educação em saúde”.

Figura 35 – Mapa de calor de 10 itens avaliando a aceitação de tecnologia de 17 usuários usando 3 construtos do Modelo de Aceitação de Tecnologia



Fonte: Elaborado pelo autor

Os resultados da avaliação qualitativa mostram que o aplicativo AIDA foi amplamente aceito pelos especialistas, enfatizando a facilidade de uso percebida. Esse componente está relacionado ao grau em que o usuário acredita que a tecnologia é fácil de usar. As demais variáveis foram estudadas principalmente para indicar uma atitude positiva em relação ao uso da tecnologia, no que diz se refere aos construtos Utilidade Percebida e Intenção Comportamental de uso. Da mesma forma, opiniões sobre a utilidade da tecnologia no trabalho de acompanhamento, demonstram a aceitação e possibilidade de utilização da tecnologia proposta para o acompanhamento de pacientes com DM2. Em relação ao contexto de resultados e discussões, na próxima seção, segue a avaliação dos resultados obtidos pelos algoritmos aprendizado de máquina, no que se refere à classificação e predição.

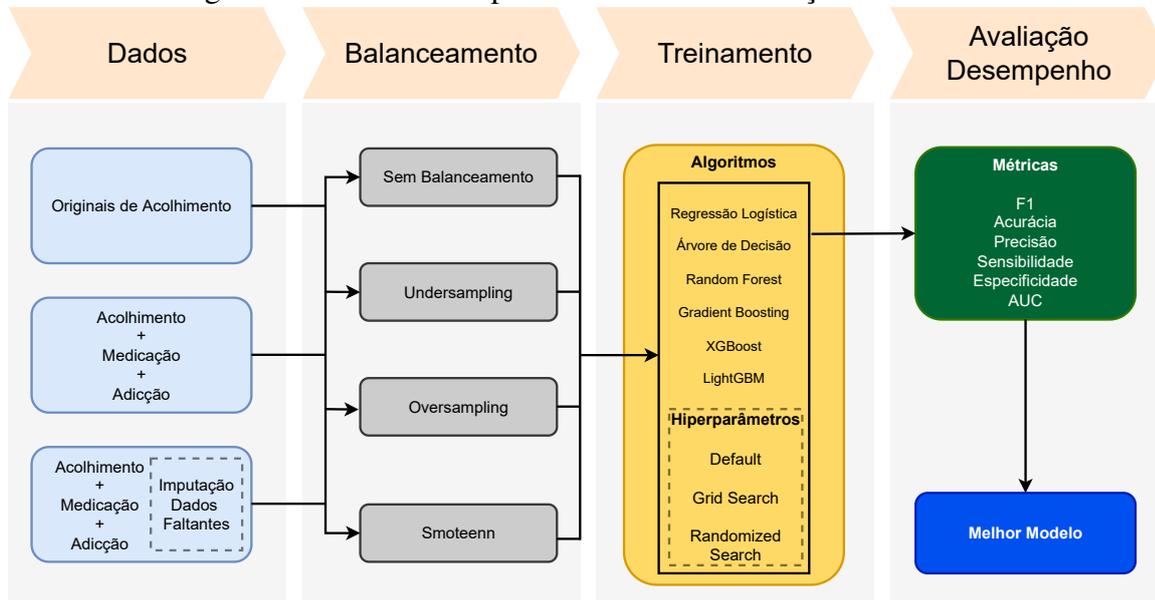
6.2 Avaliação dos algoritmos de classificação

Esta seção apresenta os resultados e discussões dos experimentos realizados para avaliar a arquitetura AIDA no que se refere ao modelo de aprendizado de máquina para classificação. A avaliação foi realizada conforme a metodologia e as métricas descritas nas Seções 5.3.2 e 5.3.4. Para avaliar cada um dos modelos gerados e produzir subsídios para comparação, foram calculadas as métricas citadas nas seções acima e conforme mostrado na Figura 36, com a descrição do objetivo de cada uma relacionada à crise de hipoglicemia em pacientes DM2:

- Matriz de confusão: esta matriz indica os erros e acertos do modelo.

- Acurácia: indica quanto o modelo acertou nas classificações gerais.
- Precisão: dos pacientes que o modelo indicou que desenvolveriam crise hipoglicêmica, quantos realmente desenvolveram.
- Sensibilidade (recall): considerando os pacientes que desenvolveram hipoglicemia, quanto o modelo acertou.
- Escore F1: é uma média harmônica entre a precisão e a sensibilidade.
- Especificidade: considerando os pacientes que não desenvolveram hipoglicemia, quanto o modelo acertou.
- AUC: consiste em uma representação gráfica da sensibilidade versus especificidade.

Figura 36 – Fluxo de Experimentos de Classificação Realizados



Fonte: Elaborado pelo autor

Os algoritmos de aprendizagem de máquina foram executados sobre três conjuntos originais de dados, os dados originais de acolhimento, dados de acolhimento com o acréscimo de dados de adicção e uso de medicamentos e, este último, com a imputação de dados ausentes. Para cada conjunto de dados foram aplicadas técnicas de balanceamento, para compensar o desbalanceamento discutido na Seção 5.2.3.3.

Foram utilizados seis algoritmos de aprendizado de máquina diferentes. Testamos a configuração inicial de cada algoritmo na sua forma padrão, sem alterar os hiperparâmetros. Na primeira sessão de testes, foram utilizados os dados brutos de acolhimento, sem imputação de dados e sem a inclusão de informações sobre vícios e medicamentos. Posteriormente, buscamos a otimização dos parâmetros utilizando as técnicas *Grid Search* e *Randomized Search*. Assim, testamos três configurações de hiperparâmetros diferentes para cada algoritmo. Da mesma forma, para cada algoritmo em diferentes configurações de hiperparâmetros, testamos com a base de dados desbalanceado e com as três opções de balanceamento (*undersampling*,

oversampling e smoteenn).

Nos resultados apresentados, apenas os melhores modelos para cada algoritmo para cada conjunto de características analisadas, são mostrados. No entanto, não há consenso sobre a melhor métrica a ser usada para avaliar modelos de aprendizado de máquina. Os médicos e desenvolvedores de software precisam entender como os modelos de ML propostos podem melhorar o atendimento ao paciente (BRUGHMANS; MELIS; MARTENS, 2023; RENGASAMY; ROTHWELL; FIGUEREDO, 2021). Nenhuma métrica única captura todas as propriedades desejáveis de um modelo, e é por isso que várias métricas são normalmente relatadas para resumir o desempenho de um modelo (HICKS et al., 2022). Para cada modelo, apresentaremos todas as métricas explicadas na Seção 5.3.4. Detalhamos a coleta de dados e funcionamento do centro clínico na Seção 5.2.1 e a construção do conjunto de dados na Seção 5.2.2. Devido ao excesso de dados ausentes no conjunto de dados, optamos por excluir dados de exames laboratoriais que faziam parte do dataset original.

O conjunto de dados resultante contém um total de 2071 registros, sendo que destes, 250 (12.07%) correspondem a casos de hipoglicemia (variável alvo) e 1821 (87.93%) são registros sem casos de hipoglicemia. Dentro dos casos de hipoglicemia há uma pequena prevalência do sexo feminino, com 137 casos (54.80%) do total, o que se repete nos casos não hipoglicêmicos (59.97%) e no total geral de pacientes, onde há uma prevalência do sexo feminino ante o masculino, na proporção de 59.34% a 40.66%, respectivamente. O conjunto de dados clínico corresponde a nove variáveis, medidas na recepção ao paciente no centro de saúde, conforme pode ser visto na Tabela 23.

Além dos dados demográficos e com origem no acolhimento ao paciente, foram acrescentadas informações que foram coletadas no campo evolução dos pacientes, conforme mostrado na Tabela 12 da Seção 5.2.2.2, como as informações relativas a adicção e o uso dos medicamentos de controle da diabetes. Ao todo, este conjunto de dados soma dezenove variáveis preditoras (características) e um resultado de saúde (variável alvo), conforme mostrado na Tabela 13, na Seção 5.2.2.3. As características demográficas e clínicas dos pacientes são mostradas na Tabela 23.

A Tabela 24 apresenta os dados de adicção (etilismo e tabagismo). Estes dados foram acrescentados a partir do campo *evolução*, do sistema de dados do centro clínico. A partir destes dados, podemos perceber uma pequena prevalência de casos de hipoglicemia em pacientes com adicção (tanto tabagismo quanto etilismo), em uma proporção de 13.60% para 11.64% no caso de etilismo e de 39.20% para 28.01% no caso do tabagismo. Os dados de adicção foram acrescentados ao conjunto de dados, pois fazem parte das recomendações da sociedade Brasileira de Diabetes como fatores de agravamento e de piora clínica da saúde geral do paciente DM2 (BERTOLUCI et al., 2022)

Por fim, a Tabela 25 apresenta os dados de utilização de medicações para controle de diabetes. Conforme as Diretrizes da Sociedade Brasileira de Diabetes (BERTOLUCI et al., 2022), determinados tipos de medicações utilizadas de forma indevida ou sem orientação, podem in-

Tabela 23 – Características demográficas e clínicas dos pacientes

Parâmetros	Crise n (%)	Não Crise n (%)	Todos
	250 (12.07%)	1821 (87.93%)	2071
Demográficos			
Idade (em anos)			
Min - Max	13.85 - 83.87	13.95 - 87.06	13.85 - 87.06
Média (D. Padrão)	58.58 (13.53)	60.57 (12.92)	60.33 (13.01)
Sexo			
Feminino	137 (54.80%)	1092 (59.97%)	1229 (59.34%)
Masculino	113 (45.20%)	729 (40.03%)	842 (40.66%)
Clínicos			
Pressão Diastólica			
Min - Max	50 - 110	44 - 127	44 - 127
Média (D. Padrão)	77.39 (11.02)	77.16 (10.59)	77.18 (10.64)
Pressão Sistólica			
Min - Max	80 - 190	72 - 240	72 - 240
Média (D. Padrão)	130.69 (16.13)	128.31 (17.34)	128.60 (17.22)
Frequência Cardíaca			
Min - Max	53 - 119	9 - 134	9 - 134
Média (D. Padrão)	82.01 (13.51)	81.05 (14.11)	81.17 (14.04)
Temperatura			
Min - Max	34.20 - 38.40	34 - 37.60	34 - 38.40
Média (D. Padrão)	36.10 (0.45)	36.08 (0.48)	36.07 (0.47)
Peso			
Min - Max	51 - 160	34 - 161	34 - 161
Média (D. Padrão)	83.93 (18.65)	85.14 (19.90)	85 (19.76)
IMC			
Min - Max	18.40 - 60.20	16.70 - 60.60	16.7 - 60.60
Média (D. Padrão)	31.82 (7.27)	32.39 (6.98)	32.32 (35.4)
Altura (cm)			
Min - Max	136 - 192	133 - 198	133 - 198
Média (D. Padrão)	162.58 (9.70)	162.13 (10.44)	162.18 (10.35)

Fonte: Elaborado pelo autor

Abreviações:

Min = Valor mínimo; Max = Valor máximo; IMC = Índice Massa Corporal.

duzir a um estado hipoglicêmico (RUAN et al., 2020). Podemos perceber com os dados da Tabela 25 que dois medicamentos se destacam como possíveis fatores indutores nos casos de hipoglicemia, a metformina com 37.20% de usuários e a insulina, com 93.60% de usuários. As demais medicações não apresentam uma proporção significativa de usuários categorizados com hipoglicemia.

Os experimentos com o conjunto de dados descritos foram divididos em etapas conforme a Figura 36, seguindo uma sequência a partir de técnicas de balanceamento, treinamento com os algoritmos de *machine learning* e modificação dos seus hiperparâmetros. Os dados utilizados nesta análise compreendem as informações de acolhimento (pré-atendimento do paciente), com dados demográficos (idade e sexo) e clínicos (Pressão diastólica e sistólica, frequência cardíaca, etc), dados de adicção (etilismo e tabagismo) e dados de uso de medicamentos (insulina, glinida, etc). As características categóricas como sexo, utilização de medicamentos e adicção

Tabela 24 – Características clínicas dos pacientes - Adicção

Parâmetros	Crise n (%)	Não Crise n (%)	Todos
	250 (12.07%)	1821 (87.93%)	2071
Etilismo			
Sim	34 (13.60%)	212 (11.64%)	246 (11.88%)
Não	216 (86.40%)	1609 (88.36%)	1825 (88.12%)
Tabagismo			
Sim	98 (39.20%)	510 (28.01%)	608 (29.36%)
Não	152 (60.80%)	1311 (71.99%)	1463 (70.64%)

Fonte: Elaborado pelo autor

foram transformadas de acordo com o método apresentado na Seção 5.2.3.2, gerando duas novas variáveis para cada variável única, como sexo (sexo_F e sexo_M) ou insulina (insulina_N e insulina_S). O objetivo é avaliar se as informações sobre sexo biológico, uso de medicações e outras variáveis categóricas influenciam a capacidade preditiva dos modelos.

Conforme mencionado na Seção 5.3.3, utilizamos as técnicas *Grid Search* e *Randomized Search* para encontrar a melhor configuração dos hiperparâmetros dos algoritmos em relação à configuração padrão da biblioteca de implementação. Como melhores resultados dessas técnicas, podemos destacar as seguintes melhorias nas métricas de avaliação dos algoritmos de aprendizado de máquina: 13,8% na AUROC do algoritmo *Decision Tree* utilizando *Grid Search*; 22,8% em *F1 score* do algoritmo *LightGBM* utilizando *Grid Search*; 2,3% em acurácia para *LightGBM* utilizando *Grid Search*; 55,8% em precisão para *Random Forest* usando *Randomized Search*; 9,1% em Sensibilidade para *XGBoost* utilizando *Randomized Search*; 7,6% em Especificidade para *Decision Tree* utilizando *Grid Search*; Essas melhorias nas métricas de desempenho justificam a importância de ajustar os hiperparâmetros dos algoritmos e denotam que, embora não haja consenso, o *Grid Search* se mostrou mais eficaz em nosso estudo.

A Tabela 26 apresenta os melhores resultados com relação aos valores da área sob a curva ROC (AUROC), destacado em cinza na coluna correspondente. Aos resultados, adicionamos a coluna Balanceamento, detalhada na Seção 5.2.3.3 para mostrar sua influência no desempenho final dos modelos. Quando apresentamos mais de uma abordagem de balanceamento para o mesmo modelo, significa que esses métodos de balanceamento geram exatamente o mesmo resultado para valores da AUROC. Com base nos resultados dos diversos testes que realizamos, podemos perceber uma melhora nas métricas de avaliação relacionadas a área sob a curva ROC, quando comparamos os resultados com e sem balanceamento. Destacamos os valores mais significativos que enfatizam a influência do balanceamento dos dados nos diferentes algoritmos testados, conforme mostrado na Tabela 26: 24,2% de ganho para *Decision Tree* utilizando balanceamento *Oversampling*; 7,5% de ganho para *Random Forest* sem balanceamento ou utilizando balanceamento *Oversampling*; 8,9% de ganho para o algoritmo *XGBoost* sem balanceamento ou com balanceamento *Oversampling*; 7,8% de ganho para o algoritmo *LightGBM* sem balanceamento ou com os balanceamentos *Oversampling* e *Smoteen*.

Na Tabela 26, destaca-se que o algoritmo *Random Forest* obteve o melhor desempenho para

Tabela 25 – Características clínicas dos pacientes - Medicamentos

Parâmetros	Crise	Não Crise	Todos
	n (%)	n (%)	
	250 (12.07%)	1821 (87.93%)	2071
Sulfonilureia			
Sim	13 (5.20%)	166 (9.12%)	179 (8.64%)
Não	237 (94.80%)	1655 (90.88%)	1892 (91.36%)
Glinida			
Sim	0 (0.0%)	5 (.027%)	5 (0.24%)
Não	250 (60.80%)	1816 (99.73%)	2066 (99.76%)
Metformina			
Sim	93 (37.20%)	519 (28.50%)	612 (29.55%)
Não	158 (62.80%)	1302 (71.50%)	1459 (70.45%)
Pioglitazona			
Sim	10 (4.0%)	43 (2.36%)	53 (2.56%)
Não	240 (96.00%)	1778 (97.64%)	2018 (97.44%)
Inibidor de dpp4			
Sim	5 (2.00%)	39 (2.14%)	44 (2.12%)
Não	245 (98.00%)	1782 (97.86%)	2027 (97.88%)
Inibidor de SGLT2			
Sim	7 (2.80%)	24 (1.31%)	31 (1.50%)
Não	243 (97.20%)	1797 (98.68%)	2040 (98.50%)
Inibidor de α-glicosidase			
Sim	0 (0.00%)	1 (0.05%)	1 (0.05%)
Não	250 (100%)	1820 (99.95%)	2070 (99.95%)
Insulina			
Sim	234 (93.60%)	741 (40.69%)	975 (47.08%)
Não	16 (6.40%)	1080 (59.31%)	1096 (52.92%)

Fonte: Elaborado pelo autor

a área sob a curva ROC, tanto sem balanceamento quanto com balanceamento Oversampling. Apesar do destaque na métrica AUROC, uma análise mais aprofundada revela que, à exceção da acurácia, as demais métricas apresentam desempenho significativamente abaixo do necessário para uma avaliação abrangente. Em particular, o *FI-score* registra uma defasagem de 37.2% em comparação com os algoritmos (Decision Tree e LightGBM) que demonstraram os melhores resultados para essa métrica. A baixa pontuação no *FI-score* pode ser explicada ao considerar a Tabela 27, pois esta métrica mantém uma relação direta com os valores de Precisão e Sensibilidade. Ambas as métricas dependem, para um desempenho otimizado, de um número substancialmente maior de Verdadeiros Positivos, o que, ao ser contrastado com a matriz de confusão apresentada na Tabela 27, evidencia uma taxa reduzida de acertos para Verdadeiros Positivos no algoritmo *Random Forest*.

A acurácia e a especificidade podem ser explicadas pelo alto grau de acerto de verdadeiros negativos (que pode ser visto na Tabela 27). Mesmo tendo valores mais expressivos para estas métricas, estes também se repetem nos demais algoritmos.

As métricas apresentadas na Tabela 26 foram baseadas na matriz de confusão mostrada na Tabela 27 calculada a partir do conjunto de dados de teste reservado para o desempenho final dos modelos. Na Tabela 27, pode-se ver destacado o resultado para o algoritmo *Random Forest*, que corresponde ao melhor resultado de área sob a curva ROC da Tabela 26. Apesar de ter

Tabela 26 – Performance dos Modelos de Aprendizado de Máquina (melhores AUROC)

Algoritmo	Balanc. ^a	AUROC ^b	F1	Acc. ^c	Prec. ^d	Sens. ^e	Espec. ^f
Regressão Logística	N	0,84	0,07	0,88	0,50	0,04	0,99
Decision Tree	O	0,82	0,43	0,77	0,30	0,72	0,77
Random Forest	N, O	0,86	0,27	0,88	0,53	0,18	0,98
Gradient Boosting	N	0,84	0,24	0,88	0,50	0,16	0,98
XGBoost	N, O	0,85	0,31	0,87	0,44	0,24	0,96
LightGBM	N, O, S	0,83	0,43	0,78	0,31	0,68	0,79

^aBalanciamento; ^bÁrea sob a curva ROC; ^cAcurácia; ^dPrecisão; ^eSensibilidade; ^fEspecificidade;

N=Sem balanceamento; O=*Oversampling*; S=*Smoteen*.

Fonte: Elaborado pelo autor

Tabela 27 – Matriz de Confusão dos algoritmos (melhores AUROC)

Algoritmo	VN ^a	FP ^b	FN ^c	VP ^d
Regressão Logística	363	2	48	2
Decision Tree	282	83	14	36
Random Forest	357	8	41	9
Gradient Boosting	357	8	42	8
XGBoost	350	15	38	12
LightGBM	290	75	16	34

^aVerdadeiro Negativo. ^bFalso Positivo. ^cFalso Negativo.

^dVerdadeiro Positivo.

Fonte: Elaborado pelo autor

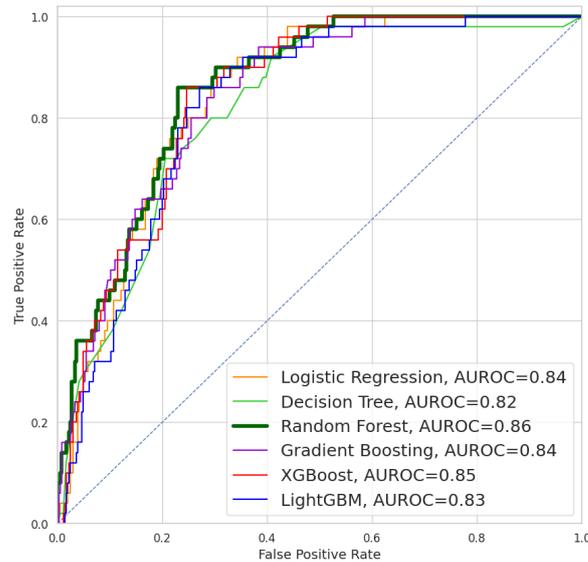
apresentado o melhor resultado sob o ponto de vista AUROC, o algoritmo *Random Forest* não apresenta os melhores resultados com relação a outras métricas, em especial ao *F1-score*. Com relação a matriz de confusão, pode-se perceber que os resultados de Verdadeiros Positivos (VP) estão muito aquém de outros algoritmos, em uma diferença que varia de nove acertos para 36 acertos no algoritmo *Decision Tree* ou ainda 34 acertos no algoritmo *LightGBM*.

A Figura 37 ilustra o gráfico da área sob a curva ROC para os seis algoritmos testados na análise realizadas com os dados de apresentados na Tabela 26. Nesta figura, destacamos o gráfico do melhor valor AUROC obtido com o algoritmo *Random Forest* com o valor de 0,86 que, apesar de ter sido o melhor resultado, os demais algoritmos não apresentaram valores substancialmente menores, tendo variado de 0,82 com o algoritmo *Decision Tree* até 0,85 do algoritmo *XGBoost*.

De acordo com a Tabela 28, o algoritmo que obteve o melhor resultado em termos de *F1-score* foi o *Decision Tree* sem balanceamento, com um valor de 0,49, mas obteve um valor de 0,72 para a AUROC. Em comparação, o algoritmo que obteve o melhor resultado em termos de AUROC (dentro do espectro de melhores *F1-score*) foi o *XGBoost* com balanceamento *Oversampling*, apresentando um valor de 0,83 mas, com um resultado de 0,44 para o *F1-score*.

As métricas apresentadas na Tabela 28 foram baseadas na matriz de confusão mostrada na Tabela 29 calculada a partir do conjunto de dados de teste reservado para o desempenho final dos modelos. Na Tabela 29, pode-se ver destacado o resultado para o algoritmo *decision Tree*, que

Figura 37 – Melhores AUROC dos Modelos



Fonte: Elaborado pelo autor

Tabela 28 – Performance dos Modelos de Aprendizado de Máquina (melhores *F1-score*)

Algoritmo	Balanc. ^a	AUROC ^b	F1	Acc. ^c	Prec. ^d	Sens. ^e	Espec. ^f
Regressão Logística	S	0,83	0,39	0,65	0,25	0,94	0,61
Decision Tree	N	0,72	0,49	0,87	0,46	0,52	0,92
Random Forest	N, S	0,82	0,44	0,75	0,30	0,80	0,75
Gradient Boosting	U	0,80	0,40	0,67	0,26	0,94	0,63
XGBoost	O	0,83	0,44	0,81	0,34	0,62	0,84
LightGBM	S	0,83	0,43	0,78	0,31	0,68	0,79

^aBalanciamento; ^bÁrea sob a curva ROC; ^cAcurácia; ^dPrecisão; ^eSensibilidade; ^fEspecificidade;

N=Sem balanceamento; O=Oversampling; S=Smoteen; U=Undersampling

Fonte: Elaborado pelo autor

corresponde ao melhor resultado de *F1-score* da Tabela 28. Apesar de ter apresentado o melhor resultado sob o ponto de vista da métrica *F1-score*, o algoritmo *Decision Tree* não apresenta os melhores resultados com relação a matriz de confusão, pode-se perceber que os resultados de Verdadeiros Positivos (VP) estão muito aquém de outros algoritmos, em uma diferença que varia de 26 acertos para 47 acertos no algoritmo Regressão Logística e no *Gradient Boosting* ou ainda 40 acertos no algoritmo *Random Forest*. Da mesma forma, as métricas para Falsos Negativos (FN) se apresenta distante de outros algoritmos, na ordem de 24 para 3 com o algoritmo de Regressão Logística ou *Gradiente Boosting*.

Os algoritmos de aprendizado de máquina utilizados neste estudo permitem investigar a importância de cada variável de entrada no resultado final apresentado. Assim, além de prever o resultado, podemos mostrar aos profissionais de saúde como os dados clínicos ou demográficos de cada paciente afetam o resultado da previsão. Na Figura 38 podemos ver a importância das variáveis de entrada de cada modelo. Na Figura 38(a), podemos visualizar as variáveis de entrada que mais afetam o resultado no modelo *Random Forest* que, conforme a Tabela 26

Tabela 29 – Matriz de Confusão dos melhores algoritmos (melhores *F1-score*)

Algoritmo	VN ^a	FP ^b	FN ^c	VP ^d
Regressão Logística	223	142	3	47
Decision Tree	334	31	24	26
Random Forest	272	93	10	40
Gradient Boosting	229	136	3	47
XGBoost	306	59	19	31
LightGBM	290	75	16	34

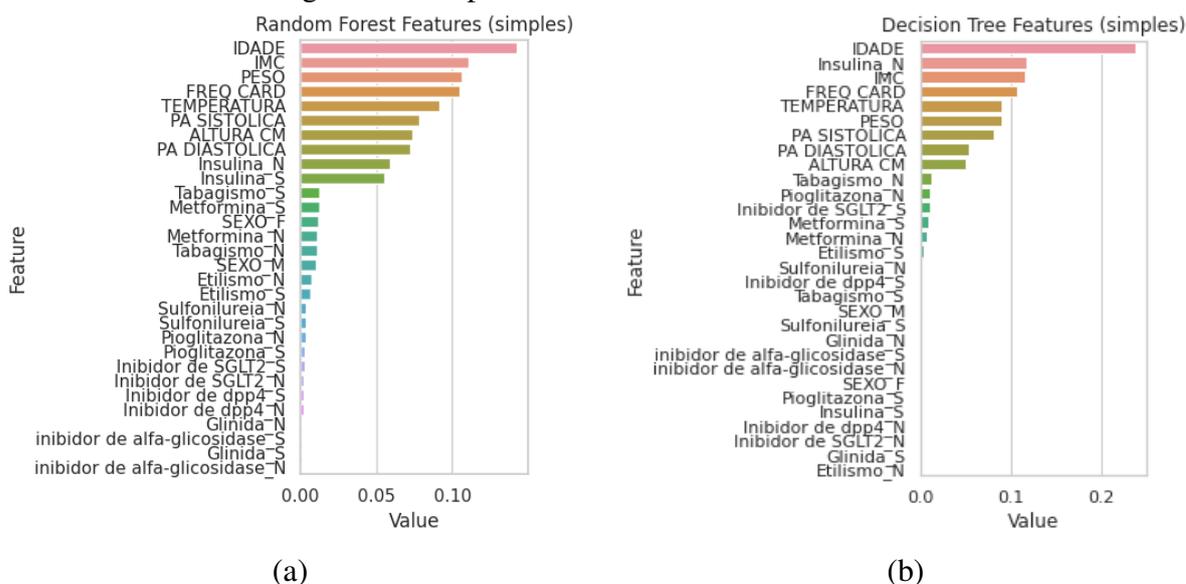
^aVerdadeiro Negativo. ^bFalso Positivo. ^cFalso Negativo.

^dVerdadeiro Positivo.

Fonte: Elaborado pelo autor

obteve o melhor resultado para a Área sob a Curva ROC. Na Figura 38(b) podemos observar as variáveis de entrada que mais influenciam o resultado no modelo *Decision Tree* que, conforme a Tabela 28 que apresentou o melhor resultado para a métrica *F1-score*.

Figura 38 – Importância das Variáveis Predictoras



Fonte: Elaborado pelo autor.

Baseado nas informações apresentadas na Figura 38 podemos observar que, em ambos os modelos, a idade do paciente figura entre as variáveis com maior importância considerando os dados de entrada. Já o sexo biológico dos pacientes não apresentou um índice de influência no modelo que utilizou o algoritmo *Random Forest*, e nenhuma influência no modelo que utilizou o algoritmo *Decision Tree* nas análises realizadas. Podemos ainda destacar a influência que as medicações utilizadas para controle do DM2 tem em ambos os modelos, com destaque para a Insulina no algoritmo *Random Forest*. Por fim, podemos destacar as variáveis de sinais vitais (Peso, Frequência Cardíaca, Pressão - sistólica e diastólica, etc) como mais significativas para os modelos apresentados na Figura 38.

Os resultados apresentados nas análises desenvolvidas, no que se refere à classificação, demonstram que os dados coletados no ambiente do Centro Clínico são importantes para a detecção de uma possível crise hipoglicêmica. A utilização dos algoritmos de aprendizado de máquina permitem avaliar um conjunto expressivo de variáveis clínicas e demográficas do paciente, o que gera uma melhor caracterização de uma crise hipoglicêmica.

6.3 Avaliação do algoritmo de Predição

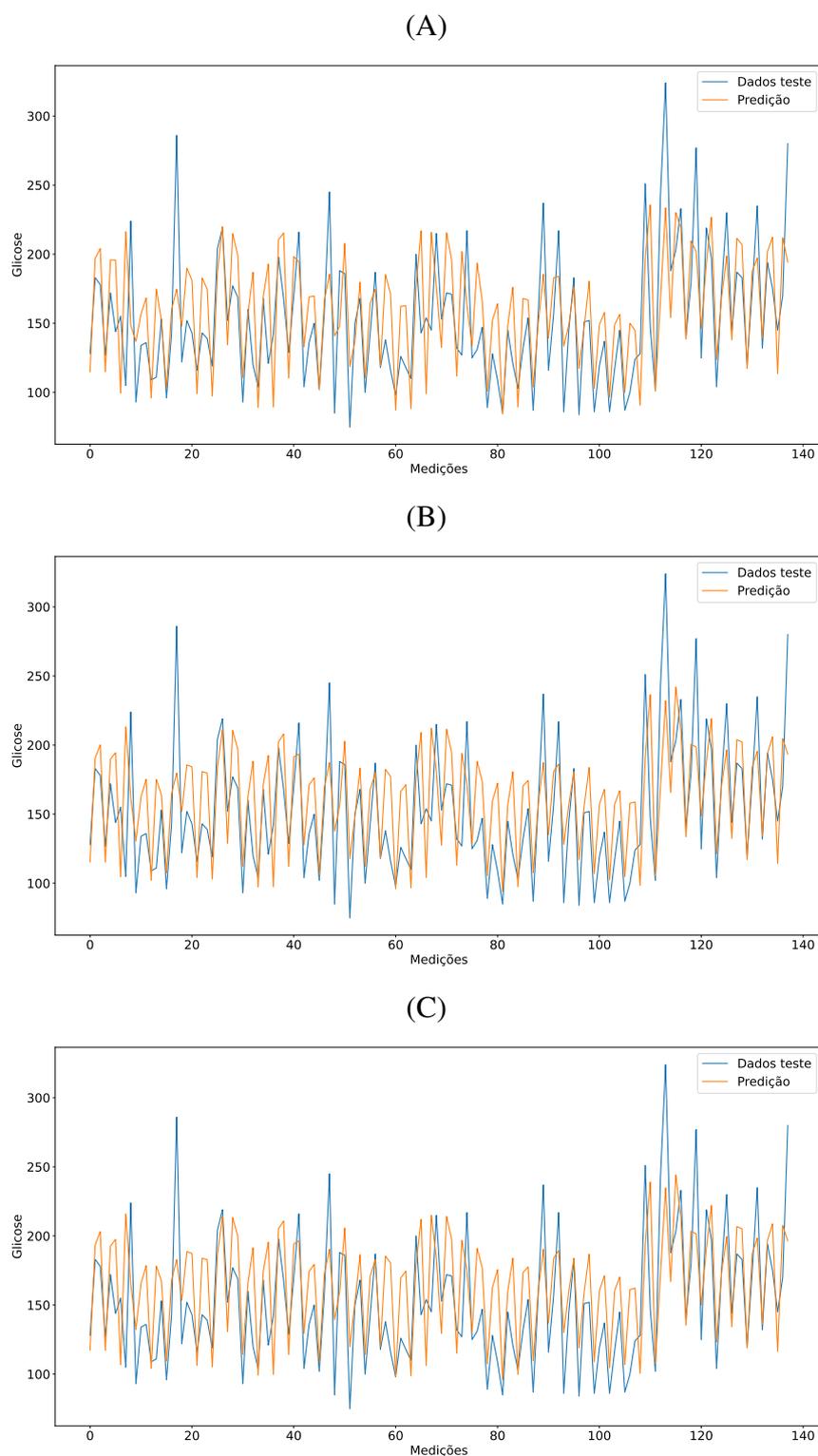
Esta seção apresenta os resultados e discussões dos experimentos realizados para avaliar a arquitetura AIDA no que se refere ao modelo de predição. A avaliação foi realizada conforme a metodologia e as métricas descritas nas Seções 5.3.2 e 5.3.4.2. A coleta de dados dos pacientes para predição ocorreu através do aplicativo AIDA descrito na seção 5.2.2.4. Este aplicativo previa a coleta dos dados do paciente da mesma forma que são coletados no acolhimento junto ao centro clínico. Assim, os dados coletados seriam os mesmos descritos com origem no acolhimento, descritos na Tabela 13.

Normalmente a literatura disponível que explora a predição de glicose em DM2 se concentra nas respostas da glicose à nutrição (ALBERS et al., 2017). No entanto, o nível de glicose é determinado por uma variedade de fatores (ALBERS et al., 2017; GIBSON et al., 2013; HORDERN et al., 2008), sendo necessário um modelo de previsão que incorpore mais variáveis, para refletir um estilo de vida em um mundo real. Dessa forma, foi feita a inclusão das variáveis preditoras descritas na Tabela 13, correspondentes ao acolhimento do paciente. O aplicativo foi utilizado por dois pacientes DM2 mas, conforme visto na Seção 5.2.3.1 houveram falhas na coleta. A primeira delas, foi que um dos pacientes simplesmente não preencheu os dados pelo aplicativo (nem de glicose e nem os demais sinais vitais). Por outro lado, o segundo paciente seguiu utilizando o aplicativo e preenchendo apenas os dados de glicose, com alguns pontos sem leitura, que podem ser visualizados na Figura 27. Desta forma, o modelo de predição foi executado sob os dados de glicose de um paciente, com coletas no período de 19/01/2022 a 05/09/2022. Os dados faltantes foram imputados seguindo a metodologia demonstrada na Seção 5.2.3.1.

Desta forma, como resultado da aplicação dos métodos de predição através do uso de redes LSTM, a Figura 39 apresenta o resultado da predição de glicose para os dados com imputação de dados faltantes conforme demonstrado nas Figuras 28(A), (B) e (C), respectivamente. Assim, a Figura 39(A) apresenta o resultado de predição para a imputação com repetição de dados imediatamente posteriores (conforme a Figura 28(A)), a Figura 39(B) apresenta o resultado de predição para imputação da média de quatro dias seguintes aos dados faltantes (conforme a Figura 28(B)). Por fim, a Figura 39(C) apresenta o resultado de predição para imputação da média de 7 dias seguintes aos dados faltantes (conforme a Figura 28(C)).

Conforme detalhado na Seção 5.3.4.2, as métricas *Root Mean Squared Error*, *Mean Absolute Error* e *Mean Poisson Deviance* foram calculadas para mensurar a qualidade da predição,

Figura 39 – Gráficos de Predição com as técnicas utilizadas para imputação de dados.



Fonte: Elaborado pelo autor

para cada tipo de imputação de dados utilizada. A Tabela 30 apresenta os resultados de cada critério de avaliação para cada tipo de imputação realizada.

Os resultados apresentados na Tabela 30 indicam que os modelos de aprendizado de má-

Tabela 30 – Resultados dos Critérios de Avaliação por Tipo de Imputação de Dados

Método de Imputação	RMSE^a	MAE^b	MPD^c
Dados Imediatamente Posteriores	39.70	32.02	9.27
Média 4 dias Posteriores	38.93	31.49	8.90
Média 7 dias Posteriores	38.74	31.41	8.82

^aRoot Mean Squared Error ^bMean Absolute Error ^cMean Poisson Deviance

Fonte: Elaborado pelo autor

quina propostos apresentaram um desempenho satisfatório na predição do nível de glicose em pacientes. Observou-se que os valores de Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE) e DESvio Médio de Poisson (MPD) foram menores nos métodos de imputação de dados que incorporaram a média dos dias subsequentes (4 e 7 dias), em comparação com o método que considerou apenas os dados imediatamente subsequentes para preencher as lacunas de dados faltantes. É perceptível que, em todos os diferentes métodos de imputação, os valores de RMSE são ligeiramente superiores aos de MAE, o que pode ser explicado pela sensibilidade do RMSE a pontos discrepantes. Os valores mais baixos para a métrica MPD resultam de sua menor penalização de valores preditos acima da média, conforme descrito na Seção 5.3.4.2. Esses resultados sugerem que a utilização de dados históricos a longo prazo pode aprimorar substancialmente melhorar a precisão das previsões. Entretanto, é importante ressaltar que esses resultados foram obtidos em um estudo específico e podem variar em diferentes contextos clínicos, tendo em vista a limitação da amostra composta por apenas um paciente.

Para avaliar o impacto clínico das previsões e oferecer suporte à tomada de decisões médicas, um critério adicional relacionado ao erro quadrático médio, denominado Erro de Grade de Clarke (*Clarke Error Grid*), foi considerado (CLARKE et al., 1987; FARUQUI et al., 2019). Isso permitiu estabelecer uma faixa aceitável para a precisão da previsão dos níveis de glicose em comparação com as observações reais. A Tabela 31 apresenta a distribuição percentual das medições em cada quadrante do Erro de Grade de Clarke para cada técnica de imputação de dados empregada, conforme detalhado na Seção 5.3.4.2.

Tabela 31 – Porcentagem de Pontos de Predição nas Zonas do Grid de Erro de Clarke.

Método de Imputação	A(%)	B(%)	C(%)	D(%)	E(%)
Dados Imediatamente Posteriores	60.14	38.41	0.00	1.45	0.00
Média 4 dias Posteriores	55.07	43.48	0.72	0.72	0.00
Média 7 dias Posteriores	55.80	42.75	0.72	0.72	0.00

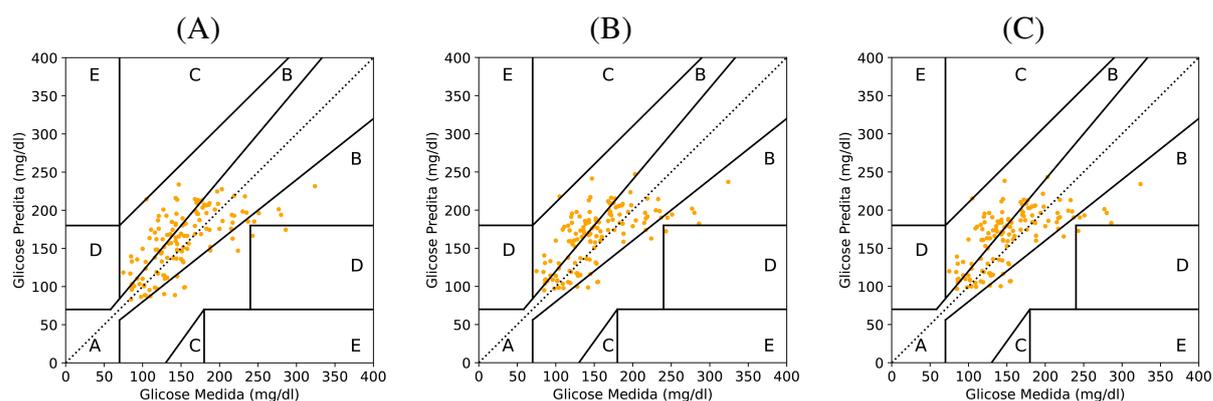
Fonte: Elaborado pelo autor

Os resultados apresentados na Tabela 30 oferecem uma visualização da distribuição percentual das previsões nas diversas zonas do Grid de Erro de Clarke para cada estratégia de imputação de dados utilizada. Notadamente, em todos os métodos de imputação, a maior proporção de pontos está concentrada na Zona A, com 60,14% para os dados imediatamente posteriores e aproximadamente 55% para os outros dois métodos (4 dias e 7 dias). Em segundo plano, as

médias de dados residem na Zona B, com mais de 40% para os métodos de média de 4 e 7 dias, e 38,41% para o método de imputação de dados imediatamente posterior. As Zonas E e D, que implicariam em tratamentos inadequados aos pacientes, apresentam percentuais nulos ou abaixo de 1,5% para todos os métodos de imputação. Até mesmo a Zona C, que aponta para tratamento inadequado sem risco substancial, somente registra porcentagens insignificantes para os métodos de média de 4 dias e 7 dias, ambas abaixo de 1%. Esses resultados convergem para um desempenho satisfatório dos modelos de previsão de glicemia, evidenciando uma sólida concordância entre as previsões e as observações reais.

Como forma de visualizar as previsões em suas respectivas zonas de Erro de Grid de Clarke, foram plotadas as previsões sob o gráfico de erros de Grid de Clarke, para os método de imputação de média dos dias imediatamente posterior (Figura 40(A)), média dos 4 dias seguintes (Figura 40(B)) e média dos 7 dias seguintes (Figura 40(C)).

Figura 40 – Erro de Grid de Clarke para cada Tipo de Imputação de Dados.



Fonte: Elaborado pelo autor

A plotagem dos dados apresentados na Figura 40 mostra que poucas previsões ficaram em Zonas de Erro de Grid de Clarke que denotam problemas no tratamento. Podemos perceber nas Figuras 40(C) e 40(B), apenas um ponto foi predito na Zona D, que conforme foi descrito na seção 5.3.4.2, pode levar a uma falha na detecção da hipoglicemia. A maioria dos pontos preditos plotados na Figura 40 estão na zona A e B que, ou estão dentro dos valores de referência para o tratamento adequado (Zona A) ou, estão fora da zona de referência mas não levam ao tratamento inadequado do paciente (Zona B).

6.4 Discussão

Nesta seção apresentamos os principais achados encontrados na pesquisa e realizamos uma discussão comparativa entre o AIDA e trabalhos representativos da literatura. É importante ressaltar que uma análise mais abrangente do estado da arte da ciência de contexto na área de saúde, com foco específico em diabetes, encontra-se detalhada no Capítulo 3. Além disso,

no Capítulo 6, apresentamos os resultados da pesquisa, dividindo-os em três principais tópicos: aceitação da tecnologia, algoritmos de classificação e algoritmos de predição. Da mesma forma, a discussão é estruturada em consonância com esses três tópicos, refletindo a organização do capítulo dos resultados.

6.4.1 Aceitação da Tecnologia

Analisando os Trabalhos Relacionados no Capítulo 3, diversos aspectos relacionados à aceitação da tecnologia na área médica são apresentados. Esta análise abrange considerações acerca da aceitação de tecnologias à luz da engenharia de usabilidade e acessibilidade, bem como a utilização da computação em nuvem para procedimentos intensivos em computação, juntamente com as preocupações subjacentes à insegurança das pessoas em relação ao compartilhamento de seus dados médicos pessoais, conforme destacado por (MICHALAKIS; CARIDAKIS, 2017). Além disso, é abordada a discussão sobre a segurança e privacidade dos dados, considerando as informações contextuais na área de saúde e seu impacto na confiança tanto de usuários como de profissionais da medicina, conforme discutido em estudos de (ZERKOUK et al., 2014; KUIJS; ROSENCRANTZ; REICH, 2015). Por outro lado, diversos artigos se concentram nos instrumentos utilizados para avaliar esses aspectos, como demonstrado em (da Silva Lima Roque et al., 2021) na validação do conteúdo e usabilidade, bem como (AHMAD et al., 2020), que investigou os fatores de influência no uso de tecnologia digital para o controle do diabetes, estendendo o modelo de avaliação TAM.

Analisando os dados da Tabela 21, que apresenta a confiabilidade do instrumento de medida avaliado por especialistas, fica evidente que os constructos demonstram uma confiabilidade substancial. Essa avaliação é sustentada pela análise dos coeficientes de Alfa de Cronbach, os quais se situaram consistentemente acima do ponto de corte estabelecido de 0.70, variando entre 0.75 e 0.80. Esses valores indicam uma alta consistência interna dos constructos. Além disso, ao examinarmos o Índice de Validade de Conteúdo por Item (I-CVI), que foi calculado para cada item de forma individual, constatamos que a maioria dos itens obteve valores superiores a 0.80, com apenas dois itens registrando pontuações abaixo desse limite. Isso confirma a confiabilidade do instrumento em nível individual para cada item dos constructos. O Índice de Validade de Conteúdo por Escala (S-CVI), que avalia a abrangência geral do roteiro do instrumento, também alcançou um valor satisfatório de 0.87. Por fim, na Figura 34, é apresentada a correlação de Pearson entre os itens dos constructos, revelando as relações lineares entre os itens dentro de seus respectivos constructos. Os resultados indicam relações positivas e significantes entre os itens, com coeficientes variando entre 0.30 e 0.75. Esses achados evidenciam que o instrumento de avaliação construído com base no Modelo de Aceitação de Tecnologia (TAM) apresenta propriedades confiáveis em sua aplicação e é adequado para avaliar as variáveis em questão.

A análise qualitativa da aceitação da tecnologia, conforme aplicação do instrumento de ava-

liação, revela uma tendência geral de concordância significativa com os aspectos relacionados à Utilidade, Facilidade de Uso e Intenção de Uso, conforme os constructos baseados no Modelo de Aceitação de Tecnologia (TAM). Entretanto, a Figura 35 evidencia que, no que diz respeito à Utilidade Percebida, ocorreram mais avaliações em que os avaliadores discordaram de certas afirmações. É importante notar que alguns avaliadores expressaram discordância ou discordância completa em relação a vários itens deste constructo, com destaque para os itens três e cinco, onde houve uma avaliação de discordância completa. Essa divergência nas avaliações pode ser atribuída a vários fatores. Primeiramente, pode estar relacionada ao fato de que, na época da aplicação do instrumento, o aplicativo ainda era um protótipo, e alguns avaliadores tiveram acesso apenas por meio de um vídeo de apresentação. Isso poderia limitar a compreensão completa da utilidade percebida do aplicativo. Além disso, a diversidade dos avaliadores, provenientes de diferentes áreas da saúde, pode ter desempenhado um papel significativo nessa divergência. Essa diversidade de perspectivas pode levar a interpretações variadas da utilidade do aplicativo, uma vez que diferentes profissionais de saúde podem ter abordagens distintas no tratamento e acompanhamento de pacientes com Diabetes Mellitus tipo 2 (DM2). Essa análise qualitativa ressalta a importância de considerar a diversidade de opiniões e contextos ao avaliar a aceitação de tecnologia em um ambiente multidisciplinar como o da saúde. Isso pode fornecer *insights* importantes para aprimorar a usabilidade e a utilidade percebida da tecnologia, especialmente quando se lida com diferentes perspectivas e experiências dos usuários.

Outra análise está relacionada a avaliação da usabilidade da tecnologia por parte dos pacientes que a utilizaram, com o objetivo de complementar as informações levantadas sobre a aceitação da tecnologia, por parte dos profissionais envolvidos no acompanhamento de pacientes DM2. Essa entrevista foi realizada conforme descrita na Seção 5.3.1.4 e Tabela 20. Assim como na avaliação pelos especialistas, a avaliação dos pacientes mostrou uma ampla aceitação da tecnologia, com opiniões como "... Achei bem tranquilo de utilizar o aplicativo.." (opinião do Paciente 1), ou ainda o Paciente 2 (quando questionado a respeito se utilizaria o app ao invés da folha de anotação): "Sim, eu usaria, porque como tenho que pegar e fazer a anotação em uma folha, não mudaria muito a lógica do que eu faço, porque eu poderia pegar e usar o aplicativo também.". Por outro lado, houveram algumas opiniões negativas a respeito do uso do aplicativo. O Paciente 1 relata que (quando indagado sobre se o App poderia ser mais fácil de utilizar): "... porque eu tenho que preencher o peso, aí salvar; preencher a glicose, aí salvar; preencher a pressão, aí salvar..", na mesma linha, o Paciente 2 relata: "... poderia ser um pouco mais simples... poderia pegar, clicar e já abrir dentro dos módulos que era pedido ..., daí tinha que pegar e abrir o outro módulo para poder pegar e colocar a funcionalidade e assim poder para simplificar, pegar e clicar e já abrir.". Essa afirmações vão ao encontro da sobrecarga de trabalho do paciente, salientada pelos especialistas, como algo que deve ser levado em consideração no desenvolvimento da tecnologia para acompanhamento dos pacientes em doenças crônicas, visando atender às necessidades e expectativas dos usuários de maneira mais eficaz.

6.4.2 Experimentos de Classificação

O Módulo de Classificação da Arquitetura AIDA objetiva realizar a detecção de pacientes em crise ou não crise de hipoglicemia, sem utilizar os valores de glicemia capilar. Nesse sentido, a avaliação dos algoritmos de machine learning com base em métricas como a Área sob a Curva ROC e o *F1-score* é uma prática bem estabelecida para determinar a eficácia desses modelos de classificação. A Tabela 26 apresenta os melhores modelos classificados de acordo com a métrica AUROC, fornecendo informações sobre quais algoritmos se destacaram na detecção de crises de hipoglicemia. A matriz de confusão, conforme demonstrada na Tabela 27, se apresenta como uma ferramenta adicional para compreender como os modelos de classificação estão se comportando em termos de verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos. Isso oferece uma compreensão mais detalhada sobre a precisão e a capacidade de detecção dos algoritmos, sob o aspecto do acompanhamento do paciente DM2. Da mesma forma, a Tabela 29 apresenta os melhores modelos com base no *F1-score*, uma métrica que considera tanto a precisão quanto a sensibilidade do modelo. Isso é especialmente útil quando se deseja equilibrar a capacidade de identificar verdadeiros positivos (VP) e minimizar falsos positivos (FP).

A análise dos resultados do modelo com a melhor avaliação, considerando o critério da curva ROC (Tabela 26), indica que o modelo alcançou um bom desempenho em termos de AUROC, atingindo um valor de 0.86, o que sugere que ele é capaz de discriminar entre as classes de pacientes em crise e não crise de hipoglicemia de maneira eficaz. Além disso, a acurácia (Acc) de 0.88 indica também um resultado positivo em termos de classificação geral. No entanto, ao examinarmos outros critérios de avaliação, como a sensibilidade (Sens), percebemos que o modelo apresentou um valor de 0.18. Isso significa que o modelo teve dificuldade em identificar corretamente os casos de pacientes em crise de hipoglicemia, o que pode ser um ponto de melhoria significativo. A precisão (Prec) do modelo, que mede quantas das classificações da classe positiva (crise de hipoglicemia) foram feitas corretamente, é de 0.53, o que indica que, das vezes em que o modelo previu uma crise, apenas 53% dessas previsões estavam corretas. O valor do F1-score de 0.27 é especialmente relevante, pois é uma média harmônica entre a precisão e a sensibilidade. Esse baixo valor sugere que o modelo está desequilibrado em relação a essas duas métricas, com um dos valores (nesse caso, a sensibilidade) muito abaixo do desejado. Esses resultados indicam que, embora o modelo tenha um desempenho sólido em termos de AUROC e acurácia geral, ele enfrenta desafios significativos na identificação precisa de pacientes em crise de hipoglicemia. Melhorar a sensibilidade do modelo é uma prioridade, pois isso permitiria identificar mais corretamente os casos de crise, o que é crucial para a aplicação clínica desse sistema de classificação. Essa análise crítica dos resultados é fundamental para orientar melhorias futuras no modelo.

A análise da matriz de confusão da Tabela 27 fornece informações importantes sobre o desempenho dos modelos de classificação, especialmente em relação ao objetivo médico de detec-

tar crises de hipoglicemia com precisão. Ao observar o modelo de classificação com o melhor desempenho de acordo com a curva ROC, o algoritmo *Random Forest* apresenta um baixo número de verdadeiros positivos (VP), o que significa que apenas 9 pacientes com hipoglicemia foram corretamente identificados como tal. Além disso, 41 pacientes com hipoglicemia foram erroneamente classificados como negativos (falsos negativos - FN). Por outro lado, ao analisar os outros modelos, como o algoritmo *Decision Tree*, que classificou 36 pacientes corretamente como verdadeiros positivos e apenas 14 como falsos negativos, ou o algoritmo *LightGBM*, que obteve 34 verdadeiros positivos e 16 falsos negativos, fica claro que esses modelos alcançaram um desempenho superior no que se refere à identificação precisa de pacientes em crise de hipoglicemia. Esses resultados são consistentes com a análise das métricas de sensibilidade, precisão e F1-score. A sensibilidade (ou taxa de verdadeiros positivos) é uma métrica crítica do ponto de vista médico, pois indica a capacidade do modelo de identificar corretamente os pacientes em crise de hipoglicemia. Os modelos *Decision Tree* e *LightGBM* superaram significativamente o *Random Forest* nessa métrica, indicando que são mais eficazes na detecção desses casos críticos. Da mesma forma, a precisão (ou taxa de verdadeiros positivos em relação a todas as classificações positivas) e o F1-score (uma média harmônica entre a precisão e a sensibilidade) refletem um desempenho mais equilibrado desses modelos em comparação ao *Random Forest*, que apresentou um desequilíbrio entre essas métricas. Portanto, do ponto de vista clínico, os algoritmos *Decision Tree* e *LightGBM* parecem ser escolhas mais adequadas para a tarefa de identificar pacientes em crise de hipoglicemia, uma vez que demonstram um desempenho superior na detecção precisa desses casos críticos.

Outra análise realizada, com base nos modelos com melhores resultados da métrica *F1-score*, conforme apresentado na Tabela 28, revela que, embora o *F1-score* tenha melhorado em relação aos melhores resultados da curva ROC, ainda não atingiu o nível desejado. O melhor modelo destacado na Tabela 28, o *Decision Tree*, possui um valor de *F1-score* de 0.49. Embora o *F1-score* seja mais equilibrado em relação à precisão e à sensibilidade do modelo, ainda assim, um valor de 0.49 indica que existe espaço para melhorias na capacidade do modelo de identificar corretamente os casos de pacientes em crise de hipoglicemia. No entanto, é importante notar que o modelo *Decision Tree* alcançou uma sensibilidade (taxa de verdadeiros positivos) de 0.52, o que é um valor positivo e indica uma melhoria em relação aos modelos com melhor desempenho na curva ROC. Essa é uma métrica crítica do ponto de vista clínico, pois reflete a capacidade do modelo de identificar corretamente os pacientes em crise de hipoglicemia. Além disso, a especificidade do modelo *Decision Tree* atingiu um valor de 0.92, o que é positivo em termos de redução de falsos positivos. No entanto, é importante observar que a área sob a curva ROC (AUROC) caiu para 0.72 em comparação com os modelos da Tabela 26. Isso sugere que, embora o modelo *Decision Tree* tenha melhorado nas métricas sensibilidade e *F1-score*, ele pode estar comprometendo a capacidade de discriminação entre classes em comparação com os modelos com melhor desempenho na curva ROC. Em resumo, a escolha entre os modelos com base na métrica *F1-score* ou na curva ROC depende dos objetivos específicos da aplicação e

das considerações clínicas. O modelo *Decision Tree* demonstra uma melhoria na identificação de casos críticos, mas pode ter uma capacidade de discriminação entre classes ligeiramente inferior em comparação com os modelos com melhor desempenho na curva ROC. Portanto, a seleção do modelo deve ser cuidadosamente ponderada em relação aos requisitos clínicos e às prioridades de classificação.

A análise dos valores da matriz de confusão para os melhores modelos com *F1-score*, conforme apresentada na Tabela 29, revela informações importantes sobre o desempenho desses modelos, especialmente sob o aspecto clínico. Observando os valores de Verdadeiros Positivos (VP), para o modelo destacado, eles são significativamente superiores aos da Tabela 27, o que indica uma melhoria na capacidade de identificar corretamente os casos de pacientes em crise de hipoglicemia. Porém, ao avaliar os demais modelos da mesma tabela, como os algoritmos de Regressão Logística e *Gradient Boosting*, que alcançaram 47 Verdadeiros Positivos e apenas 3 Falsos Negativos, fica claro que esses modelos superaram o modelo destacado em termos de identificação precisa de casos críticos. É importante notar que o aumento de Verdadeiros Positivos na Tabela 29 geralmente está acompanhado de um aumento correspondente nos Falsos Positivos. Isso sugere que os critérios de classificação desses algoritmos podem ser mais flexíveis, levando a uma maior inclusão de casos como positivos. No entanto, a manutenção dos Valores Negativos (VN) em níveis semelhantes ao modelo destacado indica que esses modelos ainda são capazes de identificar corretamente os casos negativos de maneira eficaz. Em resumo, embora o modelo destacado com base no F1-score tenha mostrado uma melhoria na identificação de casos críticos, outros modelos, como Regressão Logística e *Gradient Boosting*, demonstram um desempenho superior na identificação precisa de pacientes em crise de hipoglicemia. Essa análise da matriz de confusão fornece informações importantes sobre como os modelos estão classificando os casos e pode orientar a escolha do modelo com base em necessidades clínicas específicas.

Em resumo, a análise sob a perspectiva clínica destaca a importância de métricas como F1-score e sensibilidade, especialmente em contextos médicos onde a segurança do paciente é uma prioridade. Os modelos que alcançaram um equilíbrio favorável entre sensibilidade e precisão, como Regressão Logística e *Gradient Boosting*, podem ser considerados promissores para a detecção de crises de hipoglicemia em pacientes com DM2, mas é fundamental avaliar seu impacto clínico em termos de tratamento adequado e segurança do paciente.

Outra forma de discutir os resultados dos experimentos de classificação do trabalho, são avaliações comparativas. Porém, realizar uma comparação quantitativa direta com trabalhos relacionados se mostra desafiador devido às disparidades nos conjuntos de dados, objetivos e metodologias diversas empregadas. As variações nos conjuntos de dados e nos objetivos de classificação, aliadas à utilização de metodologias distintas, dificultam uma comparação direta. Para contornar essa complexidade, optamos por organizar uma comparação quantitativa com trabalhos de classificação semelhantes na Tabela 32, destacando aspectos específicos dos conjuntos de dados conforme nossa abordagem.

Tabela 32 – Comparação do modelo de classificação AIDA com trabalhos relacionados.

Estudo	Objetivo	Dataset	Acc ^a	Sen ^b	Spe ^c	F1 ^d	AUC ^e
Brisimi et al. (2019)	Identificar pacientes com probabilidade de serem hospitalizados durante o ano seguinte devido ao diabetes tipo II	Privado	-	-	-	-	0.89
Dworzynski et al. (2019)	Identificação de indivíduos com risco de desenvolver comorbidades decorrentes da DM2	Privado	-	-	-	-	0.80
Nanda et al. (2022)	A identificação precoce dos fatores de risco associados ao desenvolvimento da úlcera do pé diabético e classificação do tipo de úlcera	Privado	-	95.00	93.80	-	0.97
Saravanakumar e Sabibullah (2022)	Predição de bem-estar em pacientes com DM2 através de classificação	Pima-Indian	79.00	-	-	-	-
Chen et al. (2021)	Classificação do diabetes mellitus tipo 2 com ou sem comprometimento cognitivo, usando conectividade funcional de alta ordem	Privado	79.17	69.57	88.00	76.19	0.81
Pan et al. (2023)	Predição de risco para retinopatia diabética em pacientes com diabetes mellitus tipo 2 usando poucos indicadores.	Privado	79.00	00.03	-	00.06	0.70
Lee, Won e Lee (2023)	Desenvolver modelos de aprendizado de máquina que identifiquem depressão em pacientes com DM2.	Privado	90.00	69.40	-	65.10	0.83
Modelo AIDA (Auroc)	Classificação de pacientes DM2 em crise hipoglicêmica	Privado	88.00	18.00	98.00	27.00	0.86
Modelo AIDA (F1-score)	Classificação de pacientes DM2 em crise hipoglicêmica	Privado	87.00	52.00	92.00	49.00	0.72

^aAcurácia; ^bSensibilidade; ^cEspecificidade; ^dF1-score; ^eÁrea sob a curva ROC;

Fonte: Elaborado pelo autor

Na Tabela 32, podemos observar que quase a totalidade dos conjuntos de dados utilizados são privados. Esse comportamento se reflete nos objetivos de cada estudo. Por exemplo, no trabalho de (BRISIMI et al., 2019) é realizada uma previsão de hospitalização associadas ao diabetes tipo 2, no período de um ano a partir da verificação do RES do paciente, tratando a previsão como um problema de classificação e considerando a identificação de *clusters* de pacientes *a priori*. Esse tipo de classificação, com utilização de *clusters* pode melhorar o desempenho da classificação e produzir resultados mais fáceis de serem interpretados. Essas características podem ser decisivas para obter melhores métricas de desempenho quando comparadas ao nosso trabalho, como no caso da AUC.

No estudo conduzido por Dworzynski et al. (DWORZYNSKI et al., 2019), foram explorados dados de registros dinamarqueses, utilizando métodos de aprendizado de máquina comu-

mente empregados. A pesquisa teve como objetivo avaliar em que medida os registros longitudinais nacionais, abrangendo diagnósticos hospitalares, procedimentos hospitalares e prescrições ao longo de 10 anos e envolvendo quase 200.000 pacientes com Diabetes Mellitus tipo 2 (DM2), poderiam antecipar o surgimento futuro de comorbidades relacionadas a doenças crônicas nesses pacientes. A pesquisa aplicou métodos comuns de aprendizado de máquina, como regressão logística, floresta aleatória e *Gradient boosting*, nos dados de registros populacionais. Esses métodos foram empregados para prever o futuro surgimento de Infarto do Miocárdio (IM), Insuficiência Cardíaca (IC), Doença Cardiovascular (DCV), Doença Renal Crônica (DRC), Acidente Vascular Cerebral (AVC) e mortalidade por todas as causas em uma janela de cinco anos a partir do primeiro diagnóstico de DM2. Como métrica de avaliação foi utilizado a área sob curva ROC, com resultado de 0.80.

A maioria dos trabalhos comparados na Tabela 32 visa a classificação de pacientes com objetivos específicos, como a previsão de mortalidade (BRISIMI et al., 2019) ou condições de saúde mais particulares, como úlcera do pé diabético (NANDA et al., 2022) ou comprometimento cognitivo (CHEN et al., 2021). Essa especificidade contribui para a maior precisão observada nesses estudos em relação aos critérios de avaliação, em comparação com os resultados apresentados neste trabalho. A classificação e predição de condições mais suscetíveis a interpretações subjetivas, como a hipoglicemia, geralmente resultam em desempenho menos expressivo nos critérios de avaliação, conforme evidenciado nos estudos de (SARAVANAKUMAR; SABIBULLAH, 2022) e (LEE; WON; LEE, 2023), justificando resultados mais alinhados com as métricas apresentadas.

6.4.3 Experimentos de Predição

O Módulo de Regressão da Arquitetura AIDA objetiva realizar a predição de pacientes em acompanhamento, caso estejam caminhando para uma crise hipoglicêmica. Nesse sentido, para avaliar a precisão do modelo, exploramos três métricas que são bem conhecidas e amplamente utilizadas para quantificar a similaridade entre séries temporais previstas e observadas. Mais especificamente, focamos em três métricas mais frequentes na literatura de previsão do nível de glicose no sangue: *Root Mean Squared Error* RMSE, *Mean Absolute Error* (MAE) e *Mean Deviance Poisson* (MDP) (ALIBERTI et al., 2019; FARUQUI et al., 2019; TSHEPO, 2021). Embora as métricas citadas anteriormente sejam essenciais para compreender o desempenho e a precisão da previsão do modelo do ponto de vista da análise da regressão, não fornecem qualquer informação sobre o impacto clínico dos erros de previsão e das suas consequências nas decisões do tratamento médico. Então, para fornecer uma imagem mais completa do desempenho do modelo, integramos na avaliação a análise do Grid de erro de Clarke (CLARKE, 2005).

A inclusão de variáveis preditoras adicionais para prever os níveis de glicose em pacientes com DM2 é uma abordagem importante, uma vez que os níveis de glicose são influenciados por

diversos fatores além da nutrição. O estilo de vida e outros fatores de saúde podem desempenhar um papel significativo nos níveis de glicose de uma pessoa com DM2. A literatura mostra que é crucial considerar esses fatores para obter previsões mais precisas e úteis (ALBERS et al., 2017; GIBSON et al., 2013; HORDERN et al., 2008). No entanto, a coleta de dados é um desafio importante ao incorporar essas variáveis adicionais. Como mencionado na Seção 5.2.3.1, houveram falhas na coleta de dados dos pacientes que utilizaram o aplicativo AIDA. Da mesma forma, é comum encontrar dados faltantes em conjuntos de dados coletados, e a imputação de dados ausentes é uma etapa crítica para lidar com esses problemas. Porém, é importante realizar essa imputação de forma cuidadosa e considerar a natureza dos dados. Ao incorporar essas variáveis adicionais e imputar os dados ausentes, é possível construir modelos de previsão mais robustos e abrangentes que levem em consideração uma gama mais ampla de fatores que afetam os níveis de glicose em pacientes com DM2. Isso pode ser valioso para melhorar a compreensão e o gerenciamento da doença, tanto por parte dos profissionais de saúde quanto dos próprios pacientes.

Os resultados apresentados na Tabela 30 fornecem informações importantes sobre o desempenho dos modelos de aprendizado de máquina na predição do nível de glicose em pacientes com DM2, considerando os diferentes métodos de imputação de dados ausentes. Observou-se que os métodos de imputação que incorporaram a média dos dias subsequentes (4 e 7 dias) tiveram um desempenho melhor em termos de RMSE, MAE e MPD em comparação com o método que considerou apenas os dados imediatamente subsequentes. Isso sugere que a inclusão de informações de longo prazo, ou seja, uma média mais extensa, pode melhorar sensivelmente a precisão das previsões, o que pode ser confirmado no trabalho de (ALIBERTI et al., 2019). Essa descoberta é consistente com a ideia de que os níveis de glicose em pacientes com DM2 podem ser influenciados por padrões de longo prazo e tendências, além das variações de curto prazo, corroborando o que é citado em (FARUQUI et al., 2019). É importante notar que os valores de RMSE são ligeiramente superiores aos de MAE, como mencionado, devido à sensibilidade do RMSE a pontos discrepantes. No entanto, o RMSE ainda é uma métrica valiosa para avaliar o desempenho do modelo, especialmente quando se deseja penalizar erros maiores de forma mais significativa. A métrica MPD, que penaliza menos os valores preditos acima da média, também fornece informações úteis sobre a qualidade das previsões. Os valores mais baixos de MPD sugerem que os modelos estão fazendo previsões mais precisas, especialmente em relação a casos em que os valores previstos são mais altos do que a média. É importante ressaltar que esses resultados foram obtidos em um estudo específico com uma amostra composta por apenas um paciente. Essa limitação se deve à escassa adesão de participantes ao estudo. Portanto, a generalização desses resultados para diferentes contextos clínicos deve ser feita com cautela. É fundamental realizar testes em amostras maiores e em diferentes grupos de pacientes para avaliar a robustez desses modelos de predição. No entanto, esses resultados preliminares são promissores e indicam a importância de considerar métodos de imputação adequados ao lidar com dados ausentes em estudos de predição de glicose em pacientes com DM2.

Conforme demonstrado na Seção 6.3, o Grid de Erro de Clarke foi utilizado como uma forma de avaliar o impacto clínico e oferecer suporte à decisão médica na previsão de glicose. A Tabela 31 mostra a porcentagem de aceitação de cada predição de glicose nos quadrantes definidos pelo Grid de Clarke, considerando os métodos de imputação utilizados (Vide Figura 40 para referência visual). Na Tabela 31, a maior proporção de pontos de previsão está concentrada nas Zonas A e B, totalizando 98.55% dos pontos. Isso indica que a maioria das previsões está dentro de uma faixa aceitável em termos clínicos, o que está alinhado com os resultados encontrados na literatura da área (FARUQUI et al., 2019; ALIBERTI et al., 2019; Hoff, 2023). Esses resultados demonstram que o módulo de Predição da arquitetura AIDA é consistente com as descobertas de estudos anteriores, o que é um indicativo positivo de sua utilidade clínica. Essa conformidade com os resultados da literatura é importante, pois sugere que as previsões geradas pelo modelo LSTM da arquitetura AIDA têm um impacto clínico realista e podem ser utilizadas como apoio a tomada de decisões médicas no gerenciamento de pacientes com DM2.

6.5 Limitações

Quanto aos limites desta pesquisa, destacam-se os aspectos relacionados à temática "Ciência de Contexto na Saúde", que derivou da Revisão Sistemática de Literatura, bem como a incorporação da Diabetes Mellitus tipo 2 (DM2) à investigação. Além disso, os desafios envolvem a construção do *dataset* a partir de informações provenientes de um centro clínico local, a abordagem adotada nos modelos de classificação e predição e, por fim, os elementos vinculados à avaliação da aceitação e usabilidade da tecnologia aplicada ao contexto da DM2.

Em se tratando do escopo da pesquisa, esta investigação delimita-se aos aspectos relacionados exclusivamente ao Diabetes Mellitus tipo 2 (DM2), em detrimento da inclusão de outras enfermidades crônicas não transmissíveis, por exemplo. Nessa perspectiva, inicialmente, a revisão de literatura concentrou-se de maneira exclusiva em artigos que explorassem os conceitos intrínsecos à ciência de contexto, para posteriormente, direcionar-se ao DM2. Este estudo buscou responder às indagações de pesquisa delineadas a fim de obter um panorama da literatura corrente relacionada à ciência de contexto na saúde e posteriormente ao DM2, sem analisar particularmente nenhum sistema ou solução computacional específica. A pesquisa restringiu-se à obtenção de artigos publicados em distintos portais científicos associados às Tecnologias de Informação e Comunicação (TIC) e à saúde. Nossa pesquisa limitou-se aos estudos encontrados nessas plataformas durante a implementação das etapas da metodologia da revisão sistemática de literatura. Direcionamos nosso enfoque para artigos científicos, excluindo abordagens comerciais ou de natureza mais tecnológica.

Uma limitação concernente aos dados, que demanda atenção do modelo, está relacionada às informações coletadas no centro clínico, as quais, apesar de serem gerenciadas por um sistema digitalizado, contêm campos não estruturados que podem abrigar dados relevantes. Durante o processo de acolhimento ao paciente, muitos dos dados eram inicialmente registrados e posteri-

ormente revisados durante a consulta especializada. Esses últimos eram inseridos em um campo não estruturado para a correção de dados ausentes ou imprecisos. A coleta de informações nesse campo não estruturado foi realizada manualmente por bolsistas estudantes de medicina, os quais percorreram o campo de "evolução", efetuando ajustes nos dados históricos sempre que possível. Entretanto, essa correção manual pode ter introduzido novos erros de leitura e interpretação dos dados no campo não estruturado. Uma ilustração desse cenário é observada quando o campo não estruturado apresenta um valor discrepante para a hipoglicemia, indicando que o paciente estava em estado hipoglicêmico, mas sem fornecer uma data para o evento (sugerindo erroneamente que a crise hipoglicêmica estava ocorrendo no momento da consulta, o que nem sempre era verdade). Dessa maneira, é plausível que muitos dos dados relacionados à variável alvo tenham sido mal interpretados durante o processo de coleta e revisão.

Outra limitação intrínseca ao modelo está relacionada à origem e qualidade dos dados, os quais foram adquiridos de um centro clínico especializado no acompanhamento de pacientes oriundos do sistema público de saúde. Nesse contexto, os dados coletados refletem a realidade de pacientes com Diabetes Mellitus tipo 2 (DM2), caracterizados predominantemente por um baixo poder aquisitivo e baixa escolaridade. Adicionalmente, a própria natureza crônica da doença impõe a necessidade de cuidados contínuos ao longo da vida, afetando, assim, a qualidade dos dados reunidos. A responsabilidade pela coleta constante de dados, como aqueles relacionados à glicose, peso e temperatura, recai sobre o paciente. A adesão a esses cuidados e a efetivação dessa coleta emergem como desafios significativos no controle do DM2 (DIAS et al., 2023), tendo implicações diretas nos resultados dos modelos de classificação e predição.

Como decorrência da sobrecarga imposta ao paciente, a coleta de dados no âmbito do contexto do paciente foi comprometida. A adesão do paciente ao tratamento contínuo exigido pela Diabetes Mellitus tipo 2 (DM2) é um desafio amplamente documentado (MACHADO et al., 2019), sendo as ramificações desse desafio manifestadas no comprometimento do paciente em relação aos seus cuidados de longo prazo, abrangendo a medicação e a coleta de informações, principalmente relacionadas à glicose capilar. Isso resultou na escassez de dados pontuais, conforme observado no conjunto de dados do centro clínico, e na coleta de dados no contexto do paciente, realizado por meio do aplicativo AIDA. Este último evidenciou claramente tal problemática, uma vez que, após ser utilizado por dois pacientes, apenas um deles manteve a coleta de dados de forma contínua (apesar de eventuais ausências), concentrando-se exclusivamente na informação de glicose capilar, sem registrar os demais sinais vitais (peso, temperatura, etc). Por outro lado, o segundo paciente iniciou a coleta, mas desistiu após menos de um mês de utilização do aplicativo. Desta forma, os experimentos de predição do modelo AIDA foram limitados à coleta de dados de glicose capilar de um paciente.

Outra limitação está relacionada à generalização dos resultados. Os modelos de classificação foram construídos utilizando dados de pacientes com Diabetes Mellitus tipo 2 (DM2) provenientes de um centro clínico no sul do Brasil, coletados no período de 2016 a 2020. Conforme destacado por (MICHALSKI; CARBONELL; MITCHELL, 2013), a capacidade de ge-

neralização de um modelo é mais eficaz quando abrange não apenas a classificação das amostras existentes na base de treinamento, mas também a predição de um maior número de novos dados. Dessa maneira, não é possível assegurar a generalização dos resultados obtidos neste estudo para dados de pacientes de outras populações e contextos diversos. Similarmente, no que se refere ao modelo de predição, séries temporais empregando redes neurais demandam dados históricos para o treinamento da rede. Além disso, a restrição ao uso exclusivo de dados de glicemia capilar pode ser considerada uma limitação do modelo proposto, uma vez que não permite uma avaliação precisa da influência de outros sinais vitais na predição da hipoglicemia.

Por último, outra limitação da pesquisa refere-se às avaliações da aceitação e usabilidade da tecnologia. No que se refere à aceitação, a limitação presente está associada ao número restrito de amostras, o que impossibilitou uma análise estatística quantitativa, restringindo a pesquisa a uma análise qualitativa. Da mesma forma, o número reduzido de pacientes que utilizaram a tecnologia ficou limitado a dois, resultando em uma avaliação de usabilidade sujeita a ampliações futuras. No entanto, essa limitação reflete, por outro lado, as considerações discutidas na revisão bibliográfica em relação à carga de trabalho e aceitação da tecnologia por parte de pacientes com doenças crônicas (HOLZINGER; SEARLE; WERNBACHER, 2011; GUBERT; COSTA; RIGHI, 2020).

6.6 Considerações Parciais

Este capítulo apresentou os resultados obtidos dos experimentos realizados para avaliar a arquitetura AIDA. Com base nos resultados e discussões apresentados neste capítulo, é possível observar avanços significativos na avaliação da arquitetura AIDA, tanto em termos de aceitação da tecnologia pelos profissionais de saúde quanto na eficácia dos modelos de classificação e predição. A análise qualitativa dos resultados obtidos com a aplicação do instrumento de avaliação demonstraram uma robusta confiabilidade interna, fornecendo importante compreensão sobre a viabilidade e aceitação da tecnologia no contexto do acompanhamento de pacientes diabéticos tipo 2. Além disso, os experimentos de classificação e predição realizados, revelaram resultados promissores, destacando a relevância e potencial impacto clínico das decisões do modelo. No entanto, é importante ressaltar as limitações identificadas ao final do capítulo, as quais podem ser abordadas de forma a aprimorar futuras investigações e aperfeiçoar a aplicação da arquitetura AIDA no contexto da saúde.

7 CONCLUSÕES

Este estudo partiu do propósito de levantar e discutir as principais questões relativas à aplicação de ciência de contexto na saúde e identificar os conceitos de tecnologia nesta área. Para responder às questões de pesquisa deste trabalho, buscamos primeiramente sistematizar e qualificar as informações que serviram de fonte para a pesquisa. Além de responder a todas as questões específicas de pesquisa e relacioná-las na taxonomia, também podemos classificar o uso de ciência de contexto na saúde no que diz respeito aos cenários, métodos de fusão e raciocínio de contexto. As respostas e classificações obtidas contribuem para a obtenção de um grau de cobertura de pesquisas que se identificam em vários aspectos relativos à ciência de contexto na saúde. Dentro desta perspectiva, os desafios inerentes às doenças crônicas, além das questões do tratamento de doenças de longo curso aparecem como catalisadores do direcionamento da pesquisa. A dificuldade no acompanhamento do paciente e a baixa adesão ao tratamento de pacientes DM2 surgem como desafios. Além disso, o aumento dos casos de DM2 no mundo e no Brasil impõem uma sobrecarga financeira aos sistemas públicos de saúde. Uma maior aproximação do paciente com os demais *stakeholders* além da adoção de um sistema de monitoria inteligente para se antecipar ao surgimento das comorbidades surge como uma possível solução, embora ainda persistam obstáculos.

O estudo do estado da arte possibilitou a identificação de desafios e oportunidades abertas que nortearam o desenvolvimento da pesquisa. A arquitetura AIDA, proposta nesta tese, apresenta um modelo para integração dos envolvidos nos cuidados de saúde do paciente diabético tipo 2, além da integração com o registro eletrônico de saúde do paciente. Revisitando a proposta de contribuição científica, a arquitetura propõem uma subdivisão em etapas. A primeira etapa está relacionada à aquisição dos dados, o que envolve tecnologias como smartphones, aplicativos móveis, sensores, além dos dados do Registro Eletrônico de Saúde. A segunda etapa está relacionada à Inteligência Artificial para, utilizando os dados históricos do RES e dados coletados no contexto do paciente, executar a classificação do paciente em crise ou não crise hipoglicêmica. Da mesma forma, caso o paciente não esteja em crise, um módulo para realizar uma análise preditiva e prever se os dados do paciente estão induzindo a uma crise futura. Por fim, uma etapa para a apresentação dos dados aos *stakeholders*, desta forma as partes envolvidas no tratamento e acompanhamento do paciente diabético poderão ser atualizadas da situação, poderão receber alertas quando há deterioração da condição de saúde, previsão de melhoria ou piora do estado geral de saúde, além de verificar parâmetros de marcadores relacionados. Também avaliamos a aceitação da tecnologia por parte dos profissionais de saúde, como forma de verificar como a tecnologia pode ser melhorada e adaptada para a realidade dos centros de saúde e, por fim, entrevistamos pacientes que utilizaram a parte da tecnologia de coleta de dados no seu contexto e discutimos suas opiniões sob a perspectiva do *System Usability Scale* (SUS).

Este estudo descreve a implementação e a avaliação de um protótipo da arquitetura AIDA, que visa integrar registros eletrônicos de saúde com dados coletados no contexto do paciente.

O protótipo AIDA foi concebido com base em um design que promove a união dos dados históricos do paciente com informações obtidas por meio de sensores e aplicativos no ambiente do paciente. Essa integração tem o propósito de aprimorar a coleta de dados do Registro Eletrônico de Saúde (RES) e, ao fazê-lo, melhorar o tratamento de pacientes com doenças crônicas, como o Diabetes Mellitus tipo 2 (DM2). O desempenho do protótipo AIDA foi avaliado por meio de dois módulos: um para a classificação de pacientes, com o intuito de detectar crises hipoglicêmicas, e outro para a previsão de valores de glicemia. Adicionalmente, foram realizadas avaliações da aceitação da tecnologia pelos profissionais de saúde atuantes na atenção primária, bem como da usabilidade da aplicação pelos pacientes.

Este estudo foi iniciado com o propósito de oferecer uma proposta de arquitetura inovadora para apoiar a coleta de dados e monitoramento da saúde de pacientes com Diabetes Mellitus tipo 2 (DM2). Observa-se um aumento significativo no uso e armazenamento de dados de saúde dos pacientes por meio de registros eletrônicos. Entretanto, na abordagem convencional para o tratamento de pacientes com DM2, o acompanhamento ocorre em intervalos de meses, geralmente por meio de consultas na atenção primária em saúde. Nessas consultas, a possibilidade de adotar medicações para o controle da glicose é avaliada, e durante esses intervalos, o controle dos níveis de glicose e da saúde geral fica a cargo do próprio paciente. Os dados coletados pelo paciente, muitas vezes, não são incorporados ao Registro Eletrônico de Saúde (RES), sendo as informações apenas trocadas durante as consultas periódicas. Da mesma forma, durante esses intervalos, o paciente pode manifestar sintomas de alterações, como os relacionados à hipoglicemia, sem que os stakeholders tenham conhecimento ou acesso a essas informações, mesmo retrospectivamente.

Após apresentar os antecedentes essenciais para o cumprimento de nossa proposta, apresentamos o modelo AIDA. A seguir apresentamos a metodologia e trabalhos relacionados. Nesse sentido, pesquisamos trabalhos relacionados próximos à nossa proposta, que visavam enfrentar os desafios do tratamento e acompanhamento de pacientes em seu contexto. Por fim, apresentamos os resultados obtidos e os comparamos com as propostas existentes, onde podemos notar as diferenças nas propostas. Destacamos nossa proposta de arquitetura de acompanhamento de pacientes DM2. Os resultados dos módulos de classificação e previsão demonstraram a possibilidade da arquitetura ser utilizada como apoio ao serviço primário de saúde e ao tratamento e acompanhamento de pacientes DM2, demonstrando o potencial do modelo AIDA e a necessidade de maiores testes em diferentes cenários de uso.

7.1 Contribuições

Na conclusão deste trabalho, foi possível desenvolver e propor uma taxonomia abrangente para o escopo da pesquisa, resultante da análise de artigos relevantes na última década. A taxonomia identifica e agrupa diversas categorias e classificações da ciência de contexto na área da saúde, abrangendo desde a utilização de "sensores" e diferentes tipos de "modelagem" de

contexto até potenciais "cenários" de aplicação e métodos para "raciocínio de contexto". Este estudo também ampliou a revisão de literatura, incorporando o tema Diabetes Mellitus tipo 2, com o objetivo de direcionar a arquitetura para o monitoramento do tratamento de pacientes com doenças crônicas, permitindo a aplicação da taxonomia no contexto específico de pacientes com DM2. Além disso, foram identificados diversos desafios relacionados ao uso e implementação da ciência de contexto na saúde, incluindo considerações sobre a aceitação da tecnologia, tanto no que diz respeito a sensores e dispositivos vestíveis quanto à confiabilidade em sistemas médicos inteligentes. Aspectos éticos e morais também foram destacados como preocupações significativas.

A contribuição científica deste estudo reside na apresentação de uma arquitetura proposta para o gerenciamento e monitoramento da saúde de pacientes com diabetes mellitus tipo 2. A avaliação do modelo revela que a arquitetura AIDA é capaz de realizar a classificação satisfatória de pacientes entre crises e não crises hipoglicêmicas. Além disso, demonstra a capacidade de prever crises futuras com um nível de acurácia aceitável, conforme definido pela literatura médica. A aceitação da arquitetura por profissionais de saúde foi positiva, e sua utilização foi bem recebida pelos pacientes. A análise dos resultados dos algoritmos de *machine learning* no modelo de classificação mostra que os níveis de precisão em determinados algoritmos se aproximam significativamente do que seria considerado o "padrão ouro" de exames laboratoriais para determinar a presença ou ausência de crises hipoglicêmicas. De maneira análoga, o modelo de predição para glicemia, avaliado pelo padrão médico do Grid de Erro de Clarke, demonstrou um nível de precisão viável para sua finalidade, indicando que o modelo conseguiu antecipar de maneira eficaz possíveis quedas nos níveis de glicose, permitindo uma intervenção oportuna.

No que diz respeito à aceitação da tecnologia, uma investigação qualitativa preliminar foi conduzida junto a profissionais de saúde, empregando um instrumento fundamentado no *Technology Acceptance Model* (TAM), cujos resultados foram considerados satisfatórios. A maioria dos profissionais manifestou concordância quanto à percepção de facilidade de uso e à intenção de empregar a AIDA como uma ferramenta auxiliar no tratamento de pacientes com Diabetes Mellitus tipo 2 (DM2). Embora a Utilidade percebida tenha apresentado algumas discrepâncias, ainda que pequenas, essas variações podem ser atribuídas às distintas abordagens adotadas pelos profissionais, uma vez que estes provêm de diversas áreas da saúde. Finalmente, as impressões dos pacientes que utilizaram o software de coleta de dados foram avaliadas por meio de um questionário baseado na *System Usability Scale* (SUS), resultando em *feedback* positivo, com considerações favoráveis sobre o uso da tecnologia e ponderações acerca da usabilidade do software em questão.

A seguir, um resumo dos aprendizados com a condução da pesquisa:

- (a) A integração dos dados históricos dos pacientes com informações coletadas no contexto apresenta viabilidade, permitindo a criação de uma visão consolidada dos dados de saúde. Apesar de limitações, como a não execução completa do protótipo, dada sua natureza incompleta, os resultados preliminares do modelo arquitetural são

promissores;

- (b) A abordagem dos algoritmos de machine learning para detecção de crises demonstrou eficácia, indicando possibilidades de aprimoramento;
- (c) O algoritmo de predição revelou-se capaz de realizar detecção precoce dos níveis de glicemia dos pacientes, mesmo considerando uma amostragem temporal restrita;
- (d) A implementação do monitoramento para coleta de dados no contexto do paciente mostra-se como uma alternativa viável para o acompanhamento, embora seja um protótipo e careça de considerações relativas à segurança da informação;
- (e) Os resultados das avaliações de aceitação por parte dos profissionais de saúde e de usabilidade pelos pacientes indicam que há viabilidade e aceitação significativa para o uso da tecnologia no acompanhamento de pacientes em seu contexto;
- (f) Os resultados das avaliações empíricas demonstram que a arquitetura AIDA proporcionou um desempenho adequado em termos de classificação e predição de dados, com valores próximos do que seria considerado apropriado para o contexto médico.

Como contribuições adicionais, dois artigos foram publicados e disponibilizados à comunidade acadêmica, conforme lista a seguir. Um artigo foi publicado no *Universal Access in the Information Society Journal* e um artigo aceito na *9th International Conference on Internet of Things, Big Data and Security* de 2024. Além disso um artigo foi submetido no *Smart Health Journal* e se encontra em avaliação pelo periódico pós-minor review:

1. GUBERT, Luis Claudio; COSTA, Cristiano André da; RIGHI, Rodrigo da Rosa. **Context awareness in healthcare: a systematic literature review.** *Universal Access in the Information Society*, Springer, v. 19, p. 245–259, 2020.
2. GUBERT, Luis Claudio; ZEISER, Felipe A.; COSTA, Cristiano A. da; KUNST, Rafael. **Classification and Prediction of Hypoglycemia in Patients with Type 2 Diabetes Mellitus using Data from the EHR and Patient Context.** *IoTBDS*, 2024.
3. GUBERT, Luis Claudio; COSTA, Cristiano André da; KUNST, Rafael; FARIAS, Ângela Pavaglio Teixeira; JANSSEN, Luísa; BONI, Luiza. **Evaluation of an Application for Monitoring Patients with type 2 Diabetes: a Technology Acceptance Model-based qualitative analysis.** *Smart Health*, Elsevier, 2023.

7.2 Trabalhos Futuros

Em estudos futuros, vislumbramos um foco nos desafios e questões relacionadas à segurança, privacidade e confiança, que afetam diretamente a aceitação dos usuários na adoção da AIDA. Embora essas questões existam há muito tempo, ainda não possuem respostas definitivas. Outros aspectos que podem ser mais aprofundados e que são importantes para melhorar

a experiência do usuário são questões sobre usabilidade, personalização, familiaridade e conforto. Outro aspecto que pode servir como estudo futuro, com objetivo de aprofundar é explorar mais o modelo de arquitetura e implementação da AIDA acompanhando a expansão do uso de tecnologias como *wearable computing* aplicadas à saúde. Por fim, um foco na expansão da arquitetura com relação a integração com os demais stakeholders, aparece como mais um desafio.

Como trabalhos futuros, o modelo necessita de mais avaliações, principalmente em relação à segurança, privacidade e integração com outros sistemas. Além disso, como se pode verificar, existem vários desafios a serem trabalhados e respondidos, que vão desde decisões para flexibilizar o modelo em relação às regras de acesso e a integração com o registro eletrônico de saúde, até questões subjetivas que surgem. Da mesma forma, como os pacientes podem gerenciar a coleta e compartilhar seus dados na prática. Além disso, como estudo futuro, pretendemos focar no aumento das possibilidades de benefícios para pacientes e para a atenção primária em saúde. Por fim, o protótipo pode ser expandido para integrar outros padrões de dados de saúde. Outros aspectos importantes a serem discutidos estão relacionados à distribuição de dados, bem como à segurança e privacidade dos dados.

Por fim, os modelos de classificação e predição precisam sofrer testes adicionais, com um *dataset* aumentado e com mais dados de saúde, como forma de melhorar as classificações e predições. Uma questão proeminente no trabalho atual, foi a construção do *dataset*, que passou por um filtro manual realizado por profissionais de saúde, dessa forma, uma melhoria seria a completa automação da construção do *dataset*, o que iria aprimorar a eficiência da arquitetura e precisão nos resultados finais.

REFERÊNCIAS

- ABAKER, Ali A.; SAEED, Fakhreldeen A. A comparative analysis of machine learning algorithms to build a predictive model for detecting diabetes complications. **Informatica (Slovenia)**, v. 45, n. 1, p. 117–125, 2021. ISSN 18543871.
- ABDOLKHANI, Robab et al. Recommendations for the quality management of patient-generated health data in remote patient monitoring: Mixed methods study. **JMIR Mhealth Uhealth**, v. 11, p. e35917, Feb 2023. ISSN 2291-5222. Disponível em: <<https://mhealth.jmir.org/2023/1/e35917>>.
- ABHARI, Shahabeddin et al. Artificial intelligence applications in type 2 diabetes mellitus care: Focus on machine learning methods. **Healthcare Informatics Research**, v. 25, n. 4, p. 248–261, 2019. ISSN 2093369X.
- ACHARYA, U. Rajendra et al. Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. **Information Sciences**, Elsevier Inc., v. 415-416, p. 190–198, 2017. ISSN 00200255.
- AGGELIDIS, Vassilios P.; CHATZOGLOU, Prodromos D. Using a modified technology acceptance model in hospitals. **International Journal of Medical Informatics**, Elsevier, v. 78, n. 2, p. 115–126, feb 2009. ISSN 13865056.
- AHMAD, Ashfaq et al. Understanding factors influencing elderly diabetic patients' continuance intention to use digital health wearables: Extending the technology acceptance model (TAM). **Journal of Open Innovation: Technology, Market, and Complexity**, v. 6, n. 3, p. 1–15, 2020. ISSN 21998531.
- AHMADI, Hossein et al. The application of internet of things in healthcare: a systematic literature review and classification. **Universal Access in the Information Society**, Springer Berlin Heidelberg, p. 1–33, may 2018. ISSN 16155297.
- AL-NEYADI, Fahed; ABAWAJY, Jemal H. Context-based e-health system access control mechanism. In: Park, JH and Zhan, J and Lee, C and Wang, G and Kim, T and Teo, S (Ed.). **Communications in Computer and Information Science**. [S.l.: s.n.], 2009. (Communications in Computer and Information Science, v. 36), p. 68–77. ISBN 9783642026324. ISSN 18650929.
- ALANAZI, Hamdan O.; ABDULLAH, Abdul Hanan; QURESHI, Kashif Naseer. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. **Journal of Medical Systems**, Journal of Medical Systems, v. 41, n. 4, 2017. ISSN 1573689X.
- ALBERS, David J. et al. Personalized glucose forecasting for type 2 diabetes using data assimilation. **PLOS Computational Biology**, Public Library of Science, v. 13, n. 4, p. 1–38, 04 2017. Disponível em: <<https://doi.org/10.1371/journal.pcbi.1005232>>.
- ALDAHIRI, Amani; ALRASHED, Bashair; HUSSAIN, Walayat. Trends in Using IoT with Machine Learning in Health Prediction System. **Forecasting**, Multidisciplinary Digital Publishing Institute, v. 3, n. 1, p. 181–206, mar 2021. ISSN 2571-9394. Disponível em: <<https://www.mdpi.com/2571-9394/3/1/12>>.

ALIBERTI, Alessandro et al. A multi-patient data-driven approach to blood glucose prediction. **IEEE Access**, IEEE, v. 7, p. 69311–69325, 2019.

ALIREZAIE, Marjan et al. An ontology-based context-aware system for smart homes: E-care@home. **Sensors (Switzerland)**, v. 17, n. 7, p. 1–23, 2017. ISSN 14248220.

ALIZADEH, Neda et al. What we know about fatigue self-management programs for people living with chronic conditions: A scoping review. **Patient Education and Counseling**, v. 114, p. 107866, 2023. ISSN 0738-3991. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S073839912300246X>>.

ALPAYDIN, Ethem. **Introduction to Machine Learning**. 3^a ed.. ed. Cambridge, Massachusetts: MIT Press, 2014. 640 p. ISBN 9780262028189.

_____. **Machine learning : the new AI**. [S.l.: s.n.], 2016. xv, 206 pages p. ISBN 9780262529518 (pbk. alk. paper).

ALSHURAF, Nabil et al. Designing a Robust Activity Recognition Framework for Health and Exergaming Using Wearable Sensors. **IEEE Journal of Biomedical and Health Informatics**, v. 18, n. 5, p. 1636–1646, sep 2014. ISSN 21682194.

AMIN, Ruhul et al. A robust and anonymous patient monitoring system using wireless medical sensor networks. **Future Generation Computer Systems**, North-Holland, v. 80, p. 483–495, mar 2018. ISSN 0167739X.

ANDERSON, James C; GERBING, David W. Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. **Psychological Bulletin**, v. 103, n. 3, p. 411–423, 1988. ISSN 00332909.

ANDREAS, C. Müller; GUIDO, Sara. **Introduction to ML with Python**. Sebastopol, CA - USA: O'Reilly Media, 2017. 392 p. ISBN 9781449369415.

ARAÚJO, Rejane B. et al. Avaliação do cuidado prestado a pacientes diabéticos em nível primário. **Revista de Saude Publica**, v. 33, n. 1, p. 24–32, 1999. ISSN 00348910.

AZIMI, Iman et al. Internet of things for remote elderly monitoring: a study from user-centered perspective. **Journal of Ambient Intelligence and Humanized Computing**, v. 8, n. 2, p. 273–289, apr 2017. ISSN 18685145.

BAABDULLAH, Abdullah Mohammed. Consumer adoption of Mobile Social Network Games (M-SNGs) in Saudi Arabia: The role of social influence, hedonic motivation and trust. **Technology in Society**, Elsevier Ltd, v. 53, n. 2018, p. 91–102, 2018. ISSN 0160791X.

BAE, Wan D. et al. A Mobile Data Analysis Framework for Environmental Health Decision Support. In: **2012 Ninth International Conference on Information Technology - New Generations**. [S.l.: s.n.], 2012. p. 155–161. ISBN 978-1-4673-0798-7.

BAHIA, Luciana et al. Cost-effectiveness of continuous glucose monitoring with freestyle libre® in brazilian insulin-treated patients with types 1 and 2 diabetes mellitus. **Diabetology & Metabolic Syndrome**, BioMed Central, v. 15, n. 1, p. 1–12, 2023.

BAHIA, Luciana R. et al. The costs of type 2 diabetes mellitus outpatient care in the Brazilian Public Health System. **Value in Health**, Elsevier Inc., v. 14, n. 5 SUPPL., p. S137–S140, 2011. ISSN 10983015.

BALOCH, Zartasha; SHAIKH, Faisal Karim; UNAR, Mukhtiar Ali. A context-aware data fusion approach for health-iot. **International Journal of Information Technology**, v. 10, n. 3, p. 241–245, Sep 2018. ISSN 2511-2112.

BARCELO, Alberto et al. The cost of diabetes in Latin America and the Caribbean in 2015: Evidence for decision and policy makers. **Journal of Global Health**, v. 7, n. 2, 2017. ISSN 20472986.

BARROS, Aidil Jesus da Silveira; LEHFELD, Neide Aparecida de Souza. **Fundamentos de metodologia científica. 3ª edição**. [S.l.: s.n.], 2007.

BEATTIE, Mark et al. Self-management of COPD: A Technology Driven Paradigm. **Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication**, p. 53:1—53:8, 2014.

BERTOLUCI, Marcello Casaccia et al. **Diretriz Oficial da Sociedade Brasileira de Diabetes**. Conectando Pessoas, 2022. ISBN 978-65-5941-622-6. Disponível em: <<https://diretriz.diabetes.org.br>>.

BHATIA, Ridhi. Telehealth and COVID-19: Using technology to accelerate the curve on access and quality healthcare for citizens in India. **Technology in Society**, v. 64, p. 101465, feb 2021. ISSN 0160791X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0160791X20312689>>.

BICKLEY, Lynn S. **Bates' Guide to Physical Examination & History Taking**. 11th. ed. Philadelphia, US: Wolters Kluwer, 2013. 1022 p. ISBN 9781609137625.

BOBEK, Szymon; NALEPA, Grzegorz J. Uncertainty handling in rule-based mobile context-aware systems. **Pervasive and Mobile Computing**, v. 39, p. 159 – 179, 2017. ISSN 1574-1192.

Brasil, Ministério da Saúde. **Diretrizes para o cuidado das pessoas com doenças crônicas nas redes de atenção à saúde e nas linhas de cuidado prioritárias**. [S.l.: s.n.], 2013. 28 p. ISBN 9788533420182.

BRASIL. Ministério da Saúde. **RELATÓRIO DO III FÓRUM DE MONITORAMENTO DO PLANO DE AÇÕES ESTRATÉGICAS PARA O ENFRENTAMENTO DAS DOENÇAS CRÔNICAS NÃO TRANSMISSÍVEIS NO BRASIL Brasília-13 e 14 de Agosto de 2013 MINISTÉRIO DA SAÚDE Brasília-DF 2018**. [S.l.], 2018. Disponível em: <<http://portalarquivos2.saude.gov.br/images/pdf/2018/junho/19/RELATORIO-III-FORUM-DCNT-v-eletronica-13jun18-isbn.pdf>>.

BRERETON, Pearl et al. Lessons from applying the systematic literature review process within the software engineering domain. **Journal of Systems and Software**, Elsevier Inc., v. 80, n. 4, p. 571–583, 2007. ISSN 01641212.

BRISIMI, Theodora S. et al. Predicting diabetes-related hospitalizations based on electronic health records. **Statistical Methods in Medical Research**, v. 28, n. 12, p. 3667–3682, 2019. ISSN 14770334.

BRUGHMANS, Dieter; MELIS, Lissa; MARTENS, David. Disagreement amongst counterfactual explanations: How transparency can be deceptive. **arXiv preprint arXiv:2304.12667**, 2023.

BUUREN, Stef van. **Flexible Imputation of Missing Data, Second Edition**. [s.n.], 2018. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=lzb3DwAAQBAJ&oi=fnd&pg=PP1&dq=Flexible+imputation+of+missing+data&ots=Vg5WSGbb0Y&sig=YIEPsdQ03G5jy7snVYY6yDC2Af8#v=onepage&q=Flexibleimputationofmissingdata&f=false>>.

BUUREN, Stef van; GROOTHUIS-OUDSHOORN, Karin. mice: Multivariate imputation by chained equations in R. **Journal of Statistical Software**, v. 45, n. 3, p. 1–67, 2011. ISSN 15487660. Disponível em: <<https://www.jstatsoft.org/article/view/v045i03/0>>.

CARTER, Lemuria; BÉLANGER, France. The utilization of e-government services: Citizen trust, innovation and acceptance factors. **Information Systems Journal**, John Wiley & Sons, Ltd, v. 15, n. 1, p. 5–25, jan 2005. ISSN 13501917. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2575.2005.00183.x><https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2575.2005.00183.x><https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2575.2005.00183.x>>.

CENITTA, D.; ARJUNAN, R Vijaya; V, Prema K. Ischemic heart disease multiple imputation technique using machine learning algorithm. "**Engineered Science**, v. 19 (September 2022), p. 262–272, 2022. ISSN 2576-9898.

CHAWLA, Nitesh V. et al. SMOTE: Synthetic minority over-sampling technique. **Journal of Artificial Intelligence Research**, American Association for Artificial Intelligence, v. 16, p. 321–357, jun 2002. ISSN 10769757. Disponível em: <<https://www.jair.org/index.php/jair/article/view/10302>>.

CHEN, Yuna et al. Classification of type 2 diabetes mellitus with or without cognitive impairment from healthy controls using high-order functional connectivity. **Human Brain Mapping**, v. 42, n. 14, p. 4671–4684, 2021. ISSN 10970193.

CHERKAOUI, El Hadi; AGOULMINE, Nazim. Context-Aware mobility management with WiFi/3G offloading for ehealth WBANs. **2014 IEEE 16th International Conference on e-Health Networking, Applications and Services, Healthcom 2014**, p. 472–476, 2015.

CHIANG, Tzu-Chiang; LIANG, Wen-Hua. A Context-Aware Interactive Health Care System Based on Ontology and Fuzzy Inference. **Journal of Medical Systems**, v. 39, n. 9, p. 105, 2015. ISSN 0148-5598.

CHRISTOBEL, Y Angeline; SIVAPRAKASAM, P. A New Classwise k Nearest Neighbor (CKNN) Method for the Classification of Diabetes Dataset. **International Journal of Engineering and Advanced Technology**, v. 2, n. 3, p. 396–400, 2013.

CLARKE, William L. The original clarke error grid analysis (ega). **Diabetes technology & therapeutics**, Mary Ann Liebert, Inc. 2 Madison Avenue Larchmont, NY 10538 USA, v. 7, n. 5, p. 776–779, 2005.

CLARKE, William L et al. Evaluating Clinical Accuracy of Systems for Self-Monitoring of Blood Glucose. **Diabetes Care**, v. 10, n. 5, p. 622–628, 09 1987. ISSN 0149-5992. Disponível em: <<https://doi.org/10.2337/diacare.10.5.622>>.

ČOLAKOVIĆ, Alem; HADŽIALIĆ, Mesud. Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues. **Computer Networks**, Elsevier, v. 144, p. 17–39, oct 2018. ISSN 13891286.

CONTRERAS, Ivan; VEHI, Josep. Artificial intelligence for diabetes management and decision support: Literature review. **Journal of Medical Internet Research**, v. 20, n. 5, 2018. ISSN 14388871.

COSTA, Cristiano André da et al. Internet of health things: Toward intelligent vital signs monitoring in hospital wards. **Artificial Intelligence in Medicine**, v. 89, p. 61 – 69, 2018. ISSN 0933-3657.

COSTA, Cristiano Andre da; YAMIN, Adenauer Correa; GEYER, Claudio Fernando Resin. Toward a general software infrastructure for ubiquitous computing. **IEEE Pervasive Computing**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 7, n. 1, p. 64–73, jan. 2008. ISSN 1536-1268.

CUMMINGS, Peter. Missing data and multiple imputation. **JAMA Pediatrics**, American Medical Association, v. 167, n. 7, p. 656–661, jul 2013. ISSN 21686203.

da Silva Lima Roque, Geicianfran et al. Content validation and usability of a chatbot of guidelines for wound dressing. **International Journal of Medical Informatics**, Elsevier, v. 151, p. 104473, jul 2021. ISSN 18728243.

DAMTIE, Shewaneh et al. The magnitude of undiagnosed diabetes mellitus, prediabetes, and associated factors among adults living in debre tabor town, northcentral ethiopia: A community-based cross-sectional study. **Heliyon**, Elsevier, v. 9, n. 7, 2023.

DAVIS, Fred D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS Quarterly: Management Information Systems**, Management Information Systems Research Center, v. 13, n. 3, p. 319–339, 1989. ISSN 02767783.

De La Iglesia, Daniel H. et al. A context-aware indoor air quality system for sudden infant death syndrome prevention. **Sensors (Switzerland)**, Multidisciplinary Digital Publishing Institute, v. 18, n. 3, p. 757, mar 2018. ISSN 14248220.

DEEN, M. Jamal. Information and communications technologies for elderly ubiquitous health-care in a smart home. **Personal and Ubiquitous Computing**, Springer-Verlag, London, UK, UK, v. 19, n. 3, p. 573–599, 2015. ISSN 16174909.

DEFRONZO, R. A. et al. **International Textbook of Diabetes Mellitus**. 4. ed. West Sussex, UK: [s.n.], 2015. 1203 p. ISBN 9780470658611.

DEY, Anind K.; ABOWD, Gregory D.; SALBER, Daniel. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. **Human-Computer Interaction**, Taylor & Francis, v. 16, n. 2-4, p. 97–166, 2001.

DIABETES, Sociedade Brasileira de. **Diretrizes da Sociedade Brasileira de Diabetes 2019-2020**. [S.l.: s.n.], 2019. 491 p.

DIAS, Renan Italo Rodrigues et al. Doenças crônicas e a baixa adesão ao tratamento farmacológico. **Brazilian Journal of Implantology and Health Sciences**, v. 5, n. 5, p. 2926–2938, 2023.

DOBRESCU, R. Human-Centered Architecture of a Medical Cyber-Physical System. **Iwocps.Hpc.Pub.Ro**, 2015.

DRITSAS, Elias; TRIGKA, Maria. Data-Driven Machine-Learning Methods for Diabetes Risk Prediction. **Sensors**, v. 22, n. 14, 2022. ISSN 14248220.

DUARTE-RAMOS, Filipa; CABRITA, José. Using a pharmacoepidemiological approach to estimate diabetes type 2 prevalence in Portugal. **Pharmacoepidemiology and Drug Safety**, v. 15, n. 4, p. 269–274, 2006. ISSN 10538569.

DUNCAM, Bruce B. et al. **Medicina Ambulatorial: Condutas de Atenção Primária Baseadas em Evidências**. 4^a edição. ed. Porto Alegre, Brazil, Brazil: Grupo A, 2013. 1976 p. ISBN 9788536326184.

Duy Le, Tan; BEURAN, Razvan; TAN, Yasuo. Comparison of the Most Influential Missing Data Imputation Algorithms for Healthcare. **Proceedings of 2018 10th International Conference on Knowledge and Systems Engineering, KSE 2018**, IEEE, p. 247–251, 2018.

DWORZYNSKI, Piotr et al. Nationwide prediction of type 2 diabetes comorbidities. **bioRxiv**, p. 664722, 2019.

ELHELW, M. et al. An integrated multi-sensing framework for pervasive healthcare monitoring. In: **Proceedings of the 3d International ICST Conference on Pervasive Computing Technologies for Healthcare**. [S.l.: s.n.], 2009. p. 1–7. ISBN 978-963-9799-42-4.

ELMALAKI, Salma; WANNER, Lucas; SRIVASTAVA, Mani. CAreDroid. **Proceedings of the 21st Annual International Conference on Mobile Computing and Networking - MobiCom '15**, p. 386–399, 2015.

ESPOSITO, Massimo et al. A smart mobile, self-configuring, context-aware architecture for personal health monitoring. **Engineering Applications of Artificial Intelligence**, v. 67, p. 136–156, jan 2018. ISSN 09521976.

FALLAH, Mina; KALHORI, Sharareh R.Niakan. Systematic review of data mining applications in patient-centered mobile-based information systems. **Healthcare Informatics Research**, v. 23, n. 4, p. 262–270, 2017. ISSN 2093369X.

FARUQUE, Md Faisal et al. Predicting diabetes mellitus and analysing risk-factors correlation. **EAI Endorsed Transactions on Pervasive Health and Technology**, v. 5, n. 20, 2020. ISSN 24117145.

FARUQUI, Syed Hasib Akhter et al. Development of a deep learning model for dynamic forecasting of blood glucose level for type 2 diabetes mellitus: Secondary analysis of a randomized controlled trial. **JMIR mHealth and uHealth**, v. 7, n. 11, p. 1–14, 2019. ISSN 22915222.

FAWCETT, Tom. An introduction to roc analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861 – 874, 2006. ISSN 0167-8655. ROC Analysis in Pattern Recognition.

FERRARI, G. et al. The economic burden of overweight and obesity in Brazil: perspectives for the Brazilian Unified Health System. **Public Health**, W.B. Saunders, v. 207, p. 82–87, jun 2022. ISSN 14765616.

FIROUZI, Farshad et al. Internet-of-Things and big data for smarter healthcare: From device to architecture, applications and analytics. **Future Generation Computer Systems**, North-Holland, v. 78, p. 583–586, jan 2018. ISSN 0167739X.

FLORENCIO, Rêncio Bento et al. Diabetes mellitus hospitalization and mortality rate according to a national database in Brazil: a longitudinal study. **BMC Public Health**, BioMed Central Ltd, v. 21, n. 1, p. 1–7, dec 2021. ISSN 14712458. Disponível em: <<https://bmcpublihealth.biomedcentral.com/articles/10.1186/s12889-021-10438-z>>.

FORKAN, Abdur et al. BDCaM: Big Data for Context-aware Monitoring - A Personalized Knowledge Discovery Framework for Assisted Healthcare. **IEEE Transactions on Cloud Computing**, p. 1–1, 2015. ISSN 2168-7161.

GALLARDO-RINCÓN, Héctor et al. Review: Type 2 diabetes in Latin America and the Caribbean: Regional and country comparison on prevalence, trends, costs and expanded prevention. **Primary Care Diabetes**, Elsevier, v. 15, n. 2, p. 352–359, apr 2021. ISSN 18780210.

GANDHI, Khyati K; PRAJAPATI, Prof Nilesh B. Diabetes prediction using feature selection and classification. **International Journal of Advance Engineering and Research Development**, v. 1, n. 05, p. 1–7, 2014. ISSN 23486406.

GARRIDO, Juan E.; PENICHET, Víctor M. R.; LOZANO, María D. A novel context-aware system to support healthcare environments. **Universal Access in the Information Society**, Jul 2018. ISSN 1615-5297.

GERS, Felix A.; SCHMIDHUBER, Jurgen; CUMMINS, Fred. Learning to forget: Continual prediction with LSTM. **IEE Conference Publication**, v. 2, n. 470, p. 850–855, 1999. ISSN 05379989.

GIA, Tuan Nguyen et al. Fog computing in healthcare Internet of Things: A case study on ECG feature extraction. **Proceedings - 15th IEEE International Conference on Computer and Information Technology, CIT 2015, 14th IEEE International Conference on Ubiquitous Computing and Communications, IUCC 2015, 13th IEEE International Conference on Dependable, Autonomic and Secure Computing, DASC 2015 and 13th IEEE International Conference on Pervasive Intelligence and Computing, PICom 2015**, p. 356–363, 2015.

GIBSON, Bryan S et al. Development and validation of a predictive model of acute glucose response to exercise in individuals with type 2 diabetes. **Diabetology & metabolic syndrome**, BioMed Central, v. 5, n. 1, p. 1–9, 2013. Disponível em: <<https://doi.org/10.1186/1758-5996-5-33>>.

GOLDMAN, Lee; SCHAFER, Andrew i. **Goldman's Cecil Medicine**. 24. ed. Philadelphia, US: Elsevier, 2011. 2704 p. ISBN 9780808924371.

GOMES, Berto de Tácio Pereira et al. A comprehensive and scalable middleware for Ambient Assisted Living based on cloud computing and Internet of Things. **Concurrency and Computation: Practice and Experience**, v. 29, n. 11, p. e4043, 2017. ISSN 15320626.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning: Machine Learning Book**. p. 785, 2016.

GUBERT, Luis Claudio; COSTA, Cristiano André da; RIGHI, Rodrigo da Rosa. Context awareness in healthcare: a systematic literature review. **Universal Access in the Information Society**, Springer, v. 19, p. 245–259, 2020.

GUIDONI, Camilo Molino et al. Assistência ao diabetes no Sistema Único de Saúde: Análise do modelo atual. **Brazilian Journal of Pharmaceutical Sciences**, v. 45, n. 1, p. 37–48, 2009. ISSN 19848250.

HACKELING, Gavin. **Mastering machine learning with scikit-learn: apply effective learning algorithms to real-world problems using scikit-learn**. [S.l.: s.n.], 2014. 221 p. ISBN 978-1-78398-836-5.

HAMET, Pavel; TREMBLAY, Johanne. Artificial intelligence in medicine. **Metabolism: Clinical and Experimental**, Elsevier Inc., v. 69, p. S36–S40, 2017. ISSN 15328600.

HAMEURLAINE, Amina et al. Ontology and rules-based model to reason on useful contextual information for providing appropriate services in u-healthcare systems. In: CAMACHO, David et al. (Ed.). **Intelligent Distributed Computing VIII**. Cham: Springer International Publishing, 2015. p. 301–310.

Harun Al Rasyid, M. Udin et al. Anomalous data detection in WBAN measurements. **International Electronics Symposium on Knowledge Creation and Intelligent Computing, IES-KCIC 2018 - Proceedings**, p. 303–309, 2019.

HASAN, Md Kamrul et al. Diabetes prediction using ensembling of different machine learning classifiers. **IEEE Access**, v. 8, p. 76516–76531, 2020. ISSN 21693536.

HAYKIN, Simon. **Neural Networks and Learning Machines**. [S.l.: s.n.], 2008. 906 p. ISSN 14337851. ISBN 9780131471399.

HICKS, Steven A. et al. On evaluation metrics for medical applications of artificial intelligence. **Scientific Reports**, v. 12, 2022. ISSN 2045-2322. Disponível em: <<https://www.nature.com/articles/s41598-022-09954-8>>.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997. ISSN 08997667.

Hoff, A.W. **Evaluation of Multiple Machine Learning Algorithms and Glycemic Variability Indices for Predictions of Glucose Levels in Type 2 Diabetes Mellitus Based on Continuous Glucose Monitoring and Activity Data**. 2023. Disponível em: <<http://essay.utwente.nl/96264/>>.

HOI, Steven C. H. et al. Online Learning: A Comprehensive Survey. v. 1, p. 1–100, 2018.

HOLZINGER, Andreas et al. What do we need to build explainable ai systems for the medical domain? dec 2017. ISSN 14015439.

HOLZINGER, Andreas; SCHAUPP, Klaus; EDER-HALBEDL, Walter. An investigation on acceptance of ubiquitous devices for the elderly in a geriatric hospital Environment: Using the Example of Person Tracking. In: **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. v. 5105 LNCS, p. 22–29. ISBN 3540705392.

HOLZINGER, Andreas; SEARLE, Gig; WERNBACHER, Michaela. The effect of previous exposure to technology on acceptance and its importance in usability and accessibility engineering. **Universal Access in the Information Society**, Springer-Verlag, v. 10, n. 3, p. 245–260, aug 2011. ISSN 16155289.

HORDERN, Matthew D et al. Determinants of changes in blood glucose response to short-term exercise training in patients with type 2 diabetes. **Clinical Science**, Portland Press Ltd., v. 115, n. 9, p. 273–281, 2008.

Inês Schmidt, Maria; Bartholow Duncan, Bruce. O enfrentamento das doenças crônicas não transmissíveis: um desafio para a sociedade brasileira. **Epidemiologia e Serviços de Saúde**, v. 20, n. 4, p. 421–423, dec 2011. ISSN 1679-4974. Disponível em: <http://scielo.iec.pa.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742011000400001&lng=en&nrm=iso&tlng=en>.

International Diabetes Federation. **IDF Diabetes Atlas Ninth**. [S.l.: s.n.], 2019. 168 p. ISBN 9782930229874.

JAUKE, Stefanie et al. Technology Acceptance of a Machine Learning Algorithm Predicting Delirium in a Clinical Setting: a Mixed-Methods Study. **Journal of Medical Systems**, Springer, v. 45, n. 4, apr 2021. ISSN 1573689X.

JAVIDI, Hamed et al. Identification of robust deep neural network models of longitudinal clinical measurements. **npj Digital Medicine**, Springer US, v. 5, n. 1, p. 1–11, 2022. ISSN 23986352.

JEON, Jouhyun et al. Predicting Glycaemia in Type 1 Diabetes Patients: Experiments in Feature Engineering and Data Imputation. **Journal of Healthcare Informatics Research**, v. 4, n. 1, p. 71–90, 2020. ISSN 2509498X.

JOHNSON, Alistair EW et al. MIMIC-III, a freely accessible critical care database. **Scientific data**, Nature Publishing Group, v. 3, n. 1, p. 1–9, 2016.

JORDAN, JoAnne Epping. **Cuidados Inovadores para as condições crônicas: Componentes Estruturais de Ação**. [S.l.], 2003. 105 p.

JORDAN, M. I.; MITCHELL, T. M. **Machine learning: Trends, perspectives, and prospects**. 2015. 255–260 p.

JOSEPH, V. Roshan. Optimal ratio for data splitting. **Statistical Analysis and Data Mining**, John Wiley & Sons, Ltd, v. 15, n. 4, p. 531–538, aug 2022. ISSN 19321872.

JOSEPH, V. Roshan; VAKAYIL, Akhil. SPLIT: An Optimal Method for Data Splitting. **Technometrics**, American Statistical Association, v. 64, n. 2, p. 166–176, 2022. ISSN 15372723.

JUNG, Yuchae. Hybrid-aware model for senior wellness service in smart home. **Sensors (Switzerland)**, v. 17, n. 5, 2017. ISSN 14248220.

JUNWEI, Kuang et al. Dynamic prediction of cardiovascular disease using improved LSTM. **International Journal of Crowd Science**, v. 3, n. 1, p. 14–25, 2019. ISSN 2398-7294.

KALITA, Deepjyoti; MIRZA, Khalid B. Ls-grunet: glucose forecasting using deep learning for closed-loop diabetes management. In: IEEE. **2022 IEEE 7th International conference for Convergence in Technology (I2CT)**. [S.l.], 2022. p. 1–6.

KAVAKIOTIS, Ioannis et al. Machine Learning and Data Mining Methods in Diabetes Research. **Computational and Structural Biotechnology Journal**, Elsevier B.V., v. 15, p. 104–116, 2017. ISSN 20010370. Disponível em: <<http://dx.doi.org/10.1016/j.csbj.2016.12.005>>.

Khalilpour Darzi, Mohammad Rasoul; NIAKI, Seyed Taghi Akhavan; KHEDMATI, Majid. Binary classification of imbalanced datasets: The case of CoIL challenge 2000. **Expert Systems with Applications**, Pergamon, v. 128, p. 169–186, aug 2019. ISSN 09574174.

KIJSANAYOTIN, Boonchai; PANNARUNOTHAI, Supasit; SPEEDIE, Stuart M. Factors influencing health information technology adoption in Thailand's community health centers: Applying the UTAUT model. **International Journal of Medical Informatics**, v. 78, n. 6, p. 404–416, jun 2009. ISSN 13865056. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1386505608002104>>.

KITCHENHAM, Barbara. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. TR/SE-0401, p. 28, 2004. ISSN 13537776.

KOLODZIEJ, Joanna; GONZÁLEZ-VÉLEZ, Horacio; KARATZA, Helen D. High-performance modelling and simulation for big data applications. **Simulation Modelling Practice and Theory**, v. 76, p. 1–2, aug 2017. ISSN 1569190X.

KOMNAKOS, Dimitris et al. Feasibility study of a joint e-health mobile high-speed and wireless sensor system. **Proceedings of the 1st ACM international conference on Pervasive Technologies Related to Assistive Environments - PETRA '08**, p. 1, 2008.

KOPITAR, Leon et al. Early detection of type 2 diabetes mellitus using machine learning-based prediction models. **Scientific Reports**, Nature Publishing Group, v. 10, n. 1, p. 1–12, jul 2020. ISSN 20452322.

KRAEMER, Frank Alexander et al. Fog Computing in Healthcare-A Review and Discussion. **IEEE Access**, v. 5, n. 2169, p. 9206–9222, 2017. ISSN 21693536.

KUIJS, Hendrik; ROSENCRANTZ, Carina; REICH, Christoph. A Context-aware, Intelligent and Flexible Ambient Assisted Living Platform Architecture. In: Lee, YW and Westphall, CB (Ed.). **Cloud Computing**. [S.l.], 2015. (Cloud Computing-International Conference on Cloud Computing GRIDs and Virtualization, c), p. 70–76. ISBN 9781612083889. ISSN 2308-4294.

KUMAR, Vinod et al. Addressing Binary Classification over Class Imbalanced Clinical Datasets Using Computationally Intelligent Techniques. **Healthcare (Switzerland)**, Multidisciplinary Digital Publishing Institute, v. 10, n. 7, p. 1293, jul 2022. ISSN 22279032.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos da metodologia científica**. [S.l.: s.n.], 2010. 320–320 p.

LEE, Ji-Yoon; WON, Doyeon; LEE, Kiheon. Machine learning-based identification and related features of depression in patients with diabetes mellitus based on the korea national health and nutrition examination survey: A cross-sectional study. **PLOS ONE**, Public Library of Science, v. 18, n. 7, p. 1–11, 07 2023. Disponível em: <<https://doi.org/10.1371/journal.pone.0288648>>.

LI, Siqi et al. Federated and distributed learning applications for electronic health records and structured medical data: a scoping review. **Journal of the American Medical Informatics Association**, Oxford University Press (OUP), v. 30, n. 12, p. 2041–2049, ago. 2023. ISSN 1527-974X. Disponível em: <<http://dx.doi.org/10.1093/jamia/ocad170>>.

LIANG, Jun et al. Adoption of electronic health records (EHRs) in China during the past 10 years: Consecutive survey data analysis and comparison of Sino-American challenges and experiences. **Journal of Medical Internet Research**, Journal of Medical Internet Research, v. 23, n. 2, p. e24813, feb 2021. ISSN 14388871. Disponível em: <<https://www.jmir.org/2021/2/e24813>>.

LIEBERMAN, Michael A.; PEET, Alisa. **Marks' Basic Medical Biochemistry: A Clinical Approach**. 5. ed. [S.l.]: Lippincott Williams & Wilkins, 2018. 2327 p. ISBN 9781496324818.

LIU, Bin et al. **Early prediction of diabetes complications from electronic health records: A multi-task survival analysis approach**. [S.l.], 2018. 101–108 p.

LOSING, Viktor et al. Personalized online learning of whole-body motion classes using multiple inertial measurement units. **Proceedings - IEEE International Conference on Robotics and Automation**, v. 2019-May, p. 9530–9536, 2019. ISSN 10504729.

LUGER, George F. **Livro Inteligência Artificial - 6ª Edição.PDF**. [S.l.: s.n.], 2013. 614 p. ISBN 9788581435503.

MACHADO, Alencar et al. Situation-awareness as a Key for Proactive Actions in Ambient Assisted Living. In: Hammoudi, S and Maciaszek, L and Cordeiro, J and Dietz, J (Ed.). **Proceedings of the 15th International Conference on Enterprise Information Systems**. [S.l.], 2013. p. 418–426. ISBN 978-989-8565-59-4.

MACHADO, Ana Paula Morais Corrêa et al. Avaliação da adesão ao tratamento de pacientes com diabetes mellitus e seus fatores associados. **Revista Eletrônica Acervo Saúde**, n. 19, p. e565–e565, 2019.

MAHANI, Aouatef; Riad Baba Ali, Ahmed. Classification Problem in Imbalanced Datasets. In: **Recent Trends in Computational Intelligence**. [S.l.]: IntechOpen, 2020.

MALTA, Deborah Carvalho et al. Mortalidade por doenças não transmissíveis no Brasil, 1990 a 2015, segundo estimativas do estudo de carga global de doenças. **Sao Paulo Medical Journal**, v. 135, n. 3, p. 213–221, 2017. ISSN 15163180.

MALTA, Deborah Carvalho; Morais Neto, Otaliba Libânio de; Silva Junior, Jarbas Barbosa da. Apresentação do plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis no Brasil, 2011 a 2022. **Epidemiologia e Serviços de Saúde**, v. 20, n. 4, p. 425–438, 2011. ISSN 1679-4974.

MALTA, Deborah Carvalho et al. Mortalidade por doenças crônicas não transmissíveis no Brasil e suas regiões, 2000 a 2011. **Epidemiologia e Serviços de Saúde**, v. 23, n. 4, p. 599–608, 2014. ISSN 1679-4974.

MANIRUZZAMAN, Md et al. Accurate Diabetes Risk Stratification Using Machine Learning: Role of Missing Value and Outliers. **Journal of Medical Systems**, Springer New York LLC, v. 42, n. 5, may 2018. ISSN 1573689X.

MARTINS, Ana Isabel et al. European Portuguese Validation of the System Usability Scale (SUS). **Procedia Computer Science**, Elsevier Masson SAS, v. 67, n. Dsai, p. 293–300, 2015. ISSN 18770509.

MASSARO, Alessandro et al. LSTM DSS automatism and dataset optimization for diabetes prediction. **Applied Sciences (Switzerland)**, v. 9, n. 17, p. 3532, aug 2019. ISSN 20763417.

MATTHEW-MAICH, Nancy et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. **JMIR mHealth and uHealth**, v. 4, n. 2, p. e29, 2016. ISSN 2291-5222.

MCHEICK, Hamid. Modeling context aware features for pervasive computing. **Procedia Computer Science**, v. 37, p. 135–142, 2014. ISSN 18770509.

MCPHERSON, Richard A.; PINCUS, Matthew R. **Henry's Clinical Diagnosis and Management by Laboratory Methods**. 23. ed. St. Louis, Missouri: [s.n.], 2017. 1823 p. ISBN 9780323295680.

MELL, Peter; GRANCE, Timothy. The NIST-National Institute of Standards and Technology-Definition of Cloud Computing. **NIST Special Publication 800-145**, p. 7, 2011.

MICHALAKIS, Konstantinos; CARIDAKIS, George. IoT Contextual Factors on Healthcare. **Advances in Experimental Medicine and Biology**, v. 989, p. 189–200, 2017. ISSN 22148019.

MICHALSKI, Ryszard Stanislaw; CARBONELL, Jaime Guillermo; MITCHELL, Tom M. **Machine learning: An artificial intelligence approach**. [S.l.]: Springer Science & Business Media, 2013.

MIDRONI, Cooper et al. Predicting glycemia in type 1 diabetes patients: Experiments with XGBoost. **CEUR Workshop Proceedings**, v. 2148, p. 79–84, 2018. ISSN 16130073.

MILEO, Alessandra et al. A logical approach to home healthcare with intelligent sensor-network support. **Computer Journal**, v. 53, n. 8, p. 1257–1276, oct 2010. ISSN 00104620.

MITCHELL, M. et al. ContextProvider: Context awareness for medical monitoring applications. **2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society**, v. 2011, p. 5244–5247, 2011. ISSN 1557-170X (Print).

MNATSAKANYAN, Zaruhi R. et al. Distributed information fusion models for regional public health surveillance. **Information Fusion**, v. 13, n. 2, p. 129–136, apr 2012. ISSN 15662535.

MONTORI, Federico et al. Machine-to-machine wireless communication technologies for the Internet of Things: Taxonomy, comparison and open issues. **Pervasive and Mobile Computing**, Elsevier, v. 50, p. 56–81, oct 2018. ISSN 1574-1192.

MOONIAN, Oveeyen et al. Recent advances in computational tools and resources for the self-management of type 2 diabetes. **Informatics for Health and Social Care**, Taylor and Francis Ltd, v. 45, n. 1, p. 77–95, jan 2020. ISSN 17538165. Disponível em: <<https://www.tandfonline.com/action/journalInformation?journalCode=imif20>>.

MOWER, William R. et al. Pulse oximetry as a fifth vital sign in emergency geriatric assessment. **Academic Emergency Medicine**, v. 5, n. 9, p. 858–865, 1998. ISSN 10696563.

MSHALI, Haider; LEMLOUMA, Tayeb; MAGONI, Damien. Adaptive monitoring system for e-health smart homes. **Pervasive and Mobile Computing**, v. 43, p. 1–19, 2018. ISSN 15741192.

MSHALI, Haider et al. A survey on health monitoring systems for health smart homes. **International Journal of Industrial Ergonomics**, Elsevier B.V, v. 66, p. 26–56, 2018. ISSN 18728219.

MUELLER; GUIDO. **Introduction to ML with Python**. [S.l.: s.n.], 1997. 121–130 p. ISBN 9781449369415.

MUNNA, Md Tahsir Ahmed et al. Prediction model for prevalence of type-2 diabetes complications with ANN approach combining with K-fold cross validation and K-means clustering. **Lecture Notes in Networks and Systems**, v. 69, p. 1031–1045, 2020. ISSN 23673389.

MURDOCH, Travis B.; DETSKY, Allan S. The inevitable application of big data to health care. **JAMA - Journal of the American Medical Association**, v. 309, n. 13, p. 1351–1352, 2013. ISSN 00987484.

NADIMI-SHAHRAKI, Mohammad H. et al. A hybrid imputation method for multi-pattern missing data: A case study on type ii diabetes diagnosis. **Electronics (Switzerland)**, Multidisciplinary Digital Publishing Institute, v. 10, n. 24, p. 3167, dec 2021. ISSN 20799292.

NAGLIS, Michael; BHATIASEVI, Veera. Why do people use fitness tracking devices in Thailand? An integrated model approach. **Technology in Society**, Elsevier Ltd, v. 58, n. May, p. 101146, 2019. ISSN 0160791X.

NANDA, Rachita et al. Machine learning algorithm to evaluate risk factors of diabetic foot ulcers and its severity. **Medical and Biological Engineering and Computing**, Springer Berlin Heidelberg, v. 60, n. 8, p. 2349–2357, 2022. ISSN 17410444.

NATARAJAN, Thamaraiselvan; BALASUBRAMANIAN, Senthil Arasu; KASILINGAM, Dharun Lingam. The moderating role of device type and age of users on the intention to use mobile shopping applications. **Technology in Society**, Elsevier Ltd, v. 53, p. 79–90, 2018. ISSN 0160791X.

NAVA-MUÑOZ, Sandra; MORÁN, Alberto L. CANoE: A Context-Aware Notification Model to Support the Care of Older Adults in a Nursing Home. **Sensors**, v. 12, n. 12, p. 11477–11504, 2012. ISSN 1424-8220.

NEWCOMBE, Lee et al. Internet of things enabled technologies for behaviour analytics in elderly person care: a survey. In: **Proceedings - 2017 IEEE International Conference on Internet of Things, IEEE Green Computing and Communications, IEEE Cyber, Physical and Social Computing, IEEE Smart Data, iThings-GreenCom-CPSCoM-SmartData 2017**. [S.l.: s.n.], 2018. v. 2018-Janua, p. 863–870. ISBN 9781538630655.

NGUYEN, Tu Ngoc; ROKICKI, Markus. On the Predictability of non-CGM Diabetes Data for Personalized Recommendation. **ACM Reference**, v. 6, aug 2018. ISSN 16130073.

NICOLAS, Patrick R. **Scala for Machine Learning**. Birmingham, UK: Packt Publishing Ltd, 2014. 520 p. ISBN 9781783558742.

_____. **Scala for Machine Learning_ Data processing, ML algorithms, smart analytics, and more 2nd Edition**. 2^o ed.. ed. Birmingham, UK: Packt Publishing Ltd, 2017. 709 p. ISBN 9781787122383.

NILSSON, N.J. **Artificial Intelligence: a new synthesis**. [S.l.]: Morgan Kaufmann Publishers, 1998. ISBN 9781558604674.

NOORDMAN, Janneke et al. Contextual factors for the successful implementation of self-management interventions for chronic diseases: A qualitative review of reviews. **Chronic Illness**, v. 20, n. 1, p. 3–22, 2024. PMID: 36744382. Disponível em: <<https://doi.org/10.1177/17423953231153337>>.

O'DONOGHUE, John; HERBERT, John. Data Management within mHealth Environments. **Journal of Data and Information Quality**, v. 4, n. 1, p. 1–20, 2012. ISSN 19361955.

OGURTSOVA, K. et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. **Diabetes Research and Clinical Practice**, Elsevier, v. 128, p. 40–50, jun 2017. ISSN 18728227.

OGURTSOVA, Katherine et al. IDF diabetes Atlas: Global estimates of undiagnosed diabetes in adults for 2021. **Diabetes Research and Clinical Practice**, Elsevier, v. 183, p. 109118, jan 2022. ISSN 18728227.

OIKONOMOU, Evangelos K; KHERA, Rohan. Machine learning in precision diabetes care and cardiovascular risk prediction. **Cardiovascular Diabetology**, Springer, v. 22, n. 1, p. 259, 2023.

PAGIATAKIS, Catherine et al. Intelligent interaction interface for medical emergencies: Application to mobile hypoglycemia management. **Smart Health**, v. 15, n. November 2019, 2020. ISSN 23526483.

PAN, Hong et al. A risk prediction model for type 2 diabetes mellitus complicated with retinopathy based on machine learning and its application in health management. **Frontiers in Medicine**, v. 10, 2023. ISSN 2296-858X. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmed.2023.1136653>>.

PASQUIER, Thomas et al. Data provenance to audit compliance with privacy policy in the Internet of Things. **Personal and Ubiquitous Computing**, Springer London, v. 22, n. 2, p. 333–344, apr 2018. ISSN 16174909.

PEDERSEN, Alma B. et al. Missing data and multiple imputation in clinical epidemiological research. **Clinical Epidemiology**, Dove Press, v. 9, p. 157–166, mar 2017. ISSN 11791349.

PEREDA, Paula et al. Direct and Indirect Costs of Diabetes in Brazil in 2016. **Annals of Global Health**, Ubiquity Press, v. 88, n. 1, p. 1–13, mar 2022. Disponível em: <[/pmc/articles/PMC8896241/](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC8896241/)> <[https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8896241/](https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC8896241/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC8896241/)>.

PETTICREW, Mark; ROBERTS, Helen. **Systematic Reviews in the Social Sciences: A Practical Guide**. [S.l.]: Malden USA: Blackwell Pub, 2006. 336 p. ISSN 1016-9040. ISBN 9781405121101.

PINTO, Andreia et al. Recommendation Systems in the Context of Diabetes Mellitus Type 2: A Bibliometric Analysis. **Lecture Notes in Networks and Systems**, Springer Science and Business Media Deutschland GmbH, v. 468 LNNS, p. 681–691, 2022. ISSN 23673389. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-031-04826-5_67>.

POLIT, Denise F.; BECK, Cheryl Tatano. The content validity index: Are you sure you know what's being reported? Critique and recommendations. **Research in Nursing and Health**, John Wiley & Sons, Ltd, v. 29, n. 5, p. 489–497, oct 2006. ISSN 01606891.

POWERS, David Martin. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. Bioinfo Publications, 2011.

PRAHALAD, Priya et al. Equitable implementation of a precision digital health program for glucose management in individuals with newly diagnosed type 1 diabetes. **Nature Medicine**, Nature Publishing Group US New York, p. 1–9, 2024.

PRIYADARSHINI, Rojalina; BARIK, Rabindra; DUBEY, Harishchandra. DeepFog: Fog Computing-Based Deep Neural Architecture for Prediction of Stress Types, Diabetes and Hypertension Attacks. **Computation**, v. 6, n. 4, p. 62, dec 2018. ISSN 2079-3197.

QI, Jun et al. Advanced internet of things for personalised healthcare systems: A survey. **Pervasive and Mobile Computing**, v. 41, p. 132 – 149, 2017. ISSN 1574-1192.

RAHMANI, Amir M. et al. Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. **Future Generation Computer Systems**, v. 78, p. 641–658, jan 2018. ISSN 0167739X.

RAKTHANMANON, Thanawin et al. Addressing big data time series: Mining trillions of time series subsequences under dynamic time warping. **ACM Transactions on Knowledge Discovery from Data**, v. 7, n. 3, 2013. ISSN 15564681.

RAMAZI, Ramin et al. Multi-modal Predictive Models of Diabetes Progression. In: **Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics - BCB '19**. New York, New York, USA: ACM Press, 2019. v. 8, p. 253–258. ISBN 9781450366663.

RAMESH, A. N. et al. Artificial intelligence in medicine. **Annals of the Royal College of Surgeons of England**, v. 86, n. 5, p. 334–338, 2004. ISSN 00358843.

RAMEZANKHANI, Azra et al. The impact of oversampling with SMOTE on the performance of 3 classifiers in prediction of type 2 diabetes. **Medical Decision Making**, SAGE Publications Inc., v. 36, n. 1, p. 137–144, jan 2016. ISSN 1552681X. Disponível em: <https://journals.sagepub.com/doi/abs/10.1177/0272989X14560647?casa_token=H01VBcrKZXAAAAAA:ZQRjUN3zXdjk_Zy9iK8eLNeU_pospqo68iHVRuzBN7KRYFqUBlemcyOi5geIJNgL1Z-KK5eLdSyvRQ>.

RASCHKA, Sebastian et al. **Machine Learning with Pytorch and Scikit-Learn: Develop Machine Learning and Deep Learning Models with Python**. [s.n.], 2022. ISBN 9781801819312. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=SVxaEAAAQBAJ&oi=fnd&pg=PP1&dq=Machine+learning+with+pytorch+and+scikit-learn:+develop+machine+learning+and+deep+learning+models+with+python&ots=1cAOCNzGeg&sig=FzDjw0IEpDvFCuioBx5OxKK5e1Y>>.

RASCHKA, Sebastian; MIRJALILI, Vahid. **Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2**. [S.l.]: Packt Publishing Ltd, 2019.

RAULT, Tifenn et al. A survey of energy-efficient context recognition systems using wearable sensors for healthcare applications. **Pervasive and Mobile Computing**, v. 37, p. 23 – 44, 2017. ISSN 1574-1192.

REDDY, Chandan K.; AGGARWAL, Caru C. (Ed.). **Healthcare Data Analytis**. Minneapolis, Minnesota, USA: Chapman and Hall/CRC, 2015. 760 p. ISBN 9781482232127.

RENGASAMY, Divish; ROTHWELL, Benjamin C.; FIGUEREDO, Graziela P. Towards a more reliable interpretation of machine learning outputs for safety-critical systems using feature importance fusion. **Applied Sciences**, v. 11, n. 24, 2021. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/11/24/11854>>.

REY, Elizabeth et al. Personalized stress monitoring: a smartphone-enabled system for quantification of salivary cortisol. **Personal and Ubiquitous Computing**, Springer London, v. 22, n. 4, p. 867–877, aug 2018. ISSN 16174909.

RHO, Mi Jung et al. Factors influencing the acceptance of telemedicine for diabetes management. **Cluster Computing**, v. 18, n. 1, p. 321–331, mar 2015. ISSN 15737543. Disponível em: <<http://link.springer.com/10.1007/s10586-014-0356-1>>.

RIGLA, Mercedes et al. Artificial Intelligence Methodologies and Their Application to Diabetes. **Journal of Diabetes Science and Technology**, v. 12, n. 2, p. 303–310, 2018. ISSN 19322968.

RIVERA, Luis F et al. Towards Continuous Monitoring in Personalized Healthcare Through Digital Twins. **Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering**, n. November, p. 329–335, 2019.

RÖCKER, Carsten; ZIEFLE, Martina; HOLZINGER, Andreas. From computer innovation to human integration: Current trends and challenges for pervasive health technologies. In: _____. **Pervasive Health: State-of-the-art and Beyond**. London: Springer London, 2014. p. 1–17. ISBN 978-1-4471-6413-5.

RODRÍGUEZ-RODRÍGUEZ, Ignacio et al. On the possibility of predicting glycaemia ‘on the fly’ with constrained IoT devices in type 1 diabetes mellitus patients. **Sensors (Switzerland)**, v. 19, n. 20, p. 1–19, 2019. ISSN 14248220.

ROEHRS, Alex; COSTA, Cristiano André da; da Rosa Righi, Rodrigo. OmniPHR: A distributed architecture model to integrate personal health records. **Journal of Biomedical Informatics**, Elsevier Inc., v. 71, p. 70–81, 2017. ISSN 15320464.

ROEHRS, Alex et al. Personal health records: a systematic literature review. **Journal of medical Internet research**, JMIR Publications Inc., v. 19, n. 1, 2017.

ROSA, Roger et al. Estimated hospitalizations attributable to Diabetes Mellitus within the public healthcare system in Brazil from 2008 to 2010: study DIAPS 79. **Revista da Associação Médica Brasileira**, v. 60, n. 3, p. 222–230, 2014. ISSN 0104-4230.

ROSHAN, Seyed Ehsan; ASADI, Shahrokh. Improvement of Bagging performance for classification of imbalanced datasets using evolutionary multi-objective optimization. **Engineering Applications of Artificial Intelligence**, Pergamon, v. 87, p. 103319, jan 2020. ISSN 09521976.

ROSSETTI, Sarah Collins; TIASE, Victoria. The Integration of Patient-Generated Health Data to Clinical Care. Springer, Cham, p. 115–127, 2022. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-031-07696-1_6>.

RUAN, Yue et al. Predicting the risk of inpatient hypoglycemia with machine learning using electronic health records. **Diabetes Care**, v. 43, n. 7, p. 1504–1511, 2020. ISSN 19355548.

SAMPA, Masuda Begum et al. Influence of Factors on the Adoption and Use of ICT-Based eHealth Technology by Urban Corporate People. **Journal of Service Science and Management**, v. 13, n. 01, p. 1–19, 2020. ISSN 1940-9893. Disponível em: <<https://www.scirp.org/journal/doi.aspx?doi=10.4236/jssm.2020.131001>>.

SANCHEZ-PI, Nayat; CARB, Javier. An Evaluation Method for Context – Aware Systems in U-Health. **3rd International Symposium on Ambient Intelligence (ISAmI 2012)**, p. 219–226, 2012. ISSN 1867-5662.

SANNINO, G; De Pietro, G. A mobile system for real-time context-aware monitoring of patients' health and fainting. **International Journal of Data Mining and Bioinformatics**, v. 10, n. 4, p. 407–423, 2014. ISSN 17485673.

SANTOS, Danilo F.S. et al. A Standard-Based and Context-Aware Architecture for Personal Healthcare Smart Gateways. **Journal of Medical Systems**, v. 40, n. 10, 2016. ISSN 1573689X.

SARAIVA, José et al. Diabetes mellitus no Brasil: características clínicas, padrão de tratamento e custos associados ao cuidado da doença. **Jornal Brasileiro de Economia da Saúde**, v. 8, n. 2, p. 80–90, 2016. ISSN 21752095.

SARAVANAKUMAR, Venkatesh M; SABIBULLAH, M. Wellness Prediction in Diabetes Mellitus Risks Via Machine Learning Classifiers. **International Journal of Computer Science and Network Security**, v. 22, n. 4, p. 203–208, 2022. ISSN 1738-7906.

SAÚDE, Ministério da. **Vigilância de Doenças Crônicas Não Transmissíveis (DCNT)**. 2019. Disponível em: <<https://www.saude.gov.br/vigilancia-em-saude/vigilancia-de-doencas-cronicas-nao-transmissiveis-dcnt>>.

SCIKIT-LEARN. **3.1. Cross-validation: evaluating estimator performance**. 2022. 1–12 p. Disponível em: <https://scikit-learn.org/stable/modules/cross_validation.html>.

_____. **3.2 Tuning the hyper-parameters of an estimator**. 2022. 3–6 p. Disponível em: <https://scikit-learn.org/stable/modules/grid_search.html>.

SERRANO, Karina M. et al. Assessing the telemedicine acceptance for adults in Brazil. **International Journal of Health Care Quality Assurance**, v. 34, n. 1, p. 35–51, dec 2021. ISSN 09526862. Disponível em: <<https://www.emerald.com/insight/content/doi/10.1108/IJHCQA-06-2020-0098/full/html>>.

SHUJA, Mirza; MITTAL, Sonu; ZAMAN, Majid. Effective Prediction of Type II Diabetes Mellitus Using Data Mining Classifiers and SMOTE. **Springer**, p. 195–211, 2020. Disponível em: <https://link.springer.com/chapter/10.1007/978-981-15-0222-4_17>.

SMITH, Joyce; ROBERTS, Rachel. **Vital Signs for Nurses**. Oxford, UK: Blackwell Publishing Ltd, 2011. ISBN 9781405190381.

SMITH, Jack W. et al. Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. **Proceedings - Annual Symposium on Computer Applications in Medical Care**, American Medical Informatics Association, p. 261–265, nov 1988. ISSN 01954210. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/pmc/articles/PMC2245318/>>.

SNOWDON, Jane L. et al. Direct Primary Care: A New Model for Patient-Centered Care. Springer, Cham, p. 25–47, 2022. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-031-07696-1_2>.

Sociedade Brasileira de Diabetes. Posicionamento Oficial SBD nº 01/2011. p. 1–40, 2019.

SONG, Jae gu et al. Design of context-aware middleware for mobile healthcare services. **Journal of the Chinese Institute of Engineers**, Taylor & Francis, v. 35, n. 5, p. 535–545, 2012.

SOUZA-PEREIRA, Leonice et al. **Clinical decision support systems for chronic diseases: A Systematic literature review**. [S.l.]: Elsevier Ireland Ltd, 2020.

STERNE, Jonathan AC et al. Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. **BMJ (Online)**, British Medical Journal Publishing Group, v. 339, n. 7713, p. 157–160, jun 2009. ISSN 17561833.

SUN, Hong et al. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. **Diabetes Research and Clinical Practice**, Elsevier, v. 183, p. 109119, jan 2022. ISSN 18728227.

SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement Learning: An Introduction**. 2^a ed. ed. Cambridge, Massachusetts: MIT Press, 2018. 552 p. ISBN 9780262039246.

SZWARCWALD, Celia Landmann; STOPA, Sheila Rizato; MALTA, Deborah Carvalho. Situação das principais doenças crônicas não transmissíveis e dos estilos de vida da população brasileira: Pesquisa Nacional de Saúde, 2013 e 2019. **Cadernos de Saúde Pública**, FapUNIFESP (SciELO), v. 38, n. suppl 1, 2022. ISSN 0102-311X.

TANG, Yuanliang; HE, Ying. Numerical modeling of fluid and oxygen exchanges through microcirculation for the assessment of microcirculation alterations caused by type 2 diabetes. **Microvascular Research**, Elsevier, v. 117, n. January, p. 61–73, 2018. ISSN 10959319.

TAVARES, João et al. Hefestos: an intelligent system applied to ubiquitous accessibility. **Universal Access in the Information Society**, v. 15, n. 4, p. 589–607, Nov 2016. ISSN 1615-5297.

THOMAS, A.M. et al. Smart care spaces: pervasive sensing technologies for at-home care. **International Journal of Ad Hoc and Ubiquitous Computing**, v. 16, n. 4, p. 268, 2014. ISSN 1743-8225.

TOPOL, Eric J. High-performance medicine: the convergence of human and artificial intelligence. **Nature Medicine**, Springer US, v. 25, n. 1, p. 44–56, 2019. ISSN 1546170X.

TREBATOSKI, Michael et al. Methods for Leveraging a Health Information Exchange for Public Health: Lessons Learned from the NW-PHIE Experience. **Online Journal of Public Health Informatics**, v. 2, n. 2, p. 1–15, 2010. ISSN 19472579.

TSHEPO, Chris Nokeri. **Data Science Revealed: With Feature Engineering, Data Visualization, Pipeline Development, and Hyperparameter Tuning**. 1. ed. Apress Berkeley, CA, 2021. 252 p. ISBN 978-1-4842-6869-8. Disponível em: <<https://link.springer.com/book/10.1007/978-1-4842-6870-4>>.

TURING, Alan M. Computing machinery and intelligence. **Mind**, v. 59, n. 236, p. 433, 1950.

VENKATESH, Viswanath et al. User acceptance of information technology: Toward a unified view. **MIS Quarterly: Management Information Systems**, Management Information Systems Research Center, v. 27, n. 3, p. 425–478, 2003. ISSN 02767783.

VERMANI, Shalini; PURVA, Prishu. Digitization and Data Analytics in Healthcare. **Lecture Notes in Networks and Systems**, Springer Science and Business Media Deutschland GmbH, v. 317, p. 373–383, 2022. ISSN 23673389. Disponível em: <https://link.springer.com/chapter/10.1007/978-981-16-5640-8_30>.

VITABILE, Salvatore et al. Medical data processing and analysis for remote health and activities monitoring. In: _____. **High-Performance Modelling and Simulation for Big Data Applications: Selected Results of the COST Action IC1406 cHiPSet**. Cham: Springer International Publishing, 2019. p. 186–220. ISBN 978-3-030-16272-6.

WAGNER, E H. Chronic disease management: what will it take to improve care for chronic illness? **Effective clinical practice : ECP**, v. 1, n. 1, p. 2–4, 1998. ISSN 1099-8128.

WANG, Zhou; BOVIK, Alan C. Mean squared error: Lot it or leave it? A new look at signal fidelity measures. **IEEE Signal Processing Magazine**, v. 26, n. 1, p. 98–117, 2009. ISSN 10535888.

WELLS, R. H. C. et al. **CID-10: classificação estatística internacional de doenças e problemas relacionados à saúde**. [S.l.]: EDUSP, 2011.

WETZELS, Mart et al. Consume: A privacy-preserving authorisation and authentication service for connecting with health and wellbeing APIs. **Pervasive and Mobile Computing**, Elsevier, v. 43, p. 20–26, jan 2018. ISSN 15741192.

World Health Organization. **Diabetes country profiles**. [S.l.], 2016. 1 p. Disponível em: <[https://www.who.int/diabetes/country-profiles/bra\[_\]en.pdf](https://www.who.int/diabetes/country-profiles/bra[_]en.pdf)>.

YI, Haibo; NIE, Zhe. On the security of MQ cryptographic systems for constructing secure Internet of medical things. **Personal and Ubiquitous Computing**, Springer London, v. 22, n. 5-6, p. 1–7, oct 2018. ISSN 16174909.

YI, Mun Y. et al. Understanding information technology acceptance by individual professionals: Toward an integrative view. **Information and Management**, North-Holland, v. 43, n. 3, p. 350–363, apr 2006. ISSN 03787206.

YIN, Hongxu et al. DiabDeep: Pervasive Diabetes Diagnosis based on Wearable Medical Sensors and Efficient Neural Networks. **ieeexplore.ieee.org**, 2019. ISSN 21686750.

YOU, Yue et al. Application of machine learning methodology to assess the performance of DIABETIMSS program for patients with type 2 diabetes in family medicine clinics in Mexico. **BMC medical informatics and decision making**, NLM (Medline), v. 19, n. 1, p. 221, nov 2019. ISSN 14726947.

YOUNUS, Muhammad et al. Prediction model for prevalence of type-2 diabetes mellitus complications using machine learning approach. In: **Data Management and Analysis**. [S.l.]: Springer, 2020. p. 103–116.

ZARKOGIANNI, Konstantia et al. A Review of Emerging Technologies for the Management of Diabetes Mellitus. **IEEE Transactions on Biomedical Engineering**, v. 62, n. 12, p. 2735–2749, dec 2015. ISSN 0018-9294.

ZERKOUK, Meriem et al. Behavior and capability based access control model for personalized telehealthcare assistance. **Mobile Networks and Applications**, v. 19, n. 3, p. 392–403, jun 2014. ISSN 1383469X.

ZHANG, He; Ali Babar, Muhammad. Systematic reviews in software engineering: An empirical investigation. **Information and Software Technology**, Elsevier B.V., v. 55, n. 7, p. 1341–1354, 2013. ISSN 09505849.

ZHANG, Weiping et al. Medical data fusion algorithm based on Internet of things. **Personal and Ubiquitous Computing**, Springer London, v. 22, n. 5-6, p. 1–8, oct 2018. ISSN 16174909.

ZHANG, Yiyu et al. Factors influencing patients' intention to use diabetes management apps based on an extended unified theory of acceptance and use of technology model: Web-based survey. **Journal of Medical Internet Research**, v. 21, n. 8, 2019. ISSN 14388871.

ZHENG, Shuai et al. Effective Information Extraction Framework for Heterogeneous Clinical Reports Using Online Machine Learning and Controlled Vocabularies. **JMIR Medical Informatics**, v. 5, n. 2, p. e12, 2017. ISSN 2291-9694.

ZIEFLE, Martina; ROCKER, Carsten; HOLZINGER, Andreas. Medical Technology in Smart Homes: Exploring the User's Perspective on Privacy, Intimacy and Trust. In: **2011 IEEE 35th Annual Computer Software and Applications Conference Workshops**. [S.l.]: IEEE, 2011. p. 410–415. ISBN 978-1-4577-0980-7.

ZOU, Quan et al. Predicting Diabetes Mellitus With Machine Learning Techniques. **Frontiers in Genetics**, v. 9, 2018. ISSN 1664-8021.