



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA

ELVERSON SOARES DE MELO

**DESVENDANDO BIOMARCADORES: UMA ANÁLISE DE ALGORITMOS DE
APRENDIZAGEM DE MÁQUINA NA DISTINÇÃO ENTRE INDIVÍDUOS
SAUDÁVEIS E PRÉ-DIABÉTICOS A PARTIR DE DADOS DE DISPOSITIVOS
VESTÍVEIS.**

RECIFE

2023

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
BACHARELADO EM SISTEMAS DE INFORMAÇÃO

ELVERSON SOARES DE MELO

**DESVENDANDO BIOMARCADORES: UMA ANÁLISE DE ALGORITMOS DE
APRENDIZAGEM DE MÁQUINA NA DISTINÇÃO ENTRE INDIVÍDUOS
SAUDÁVEIS E PRÉ-DIABÉTICOS A PARTIR DE DADOS DE DISPOSITIVOS
VESTÍVEIS.**

TCC apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de bacharel em Sistemas de Informação.

Orientadora: Maíra Araújo de Santana

RECIFE

2023

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Melo, Elverson Soares de.

Desvendando biomarcadores: uma análise de algoritmos de aprendizagem de máquina na distinção entre indivíduos saudáveis e pré-diabéticos a partir de dados de dispositivos vestíveis. / Elverson Soares de Melo. - Recife, 2023.

75 : il., tab.

Orientador(a): Maíra Araújo de Santana

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Informática, Sistemas de Informação - Bacharelado, 2023.

1. Inteligência Artificial. 2. Diabetes Mellitus. 3. Estado Pré-Diabético. 4. Dispositivos Eletrônicos Vestíveis. I. Santana, Maíra Araújo de. (Orientação). II. Título.

000 CDD (22.ed.)

ELVERSON SOARES DE MELO

DESVENDANDO BIOMARCADORES: UMA ANÁLISE DE ALGORITMOS DE APRENDIZAGEM DE MÁQUINA NA DISTINÇÃO ENTRE INDIVÍDUOS SAUDÁVEIS E PRÉ-DIABÉTICOS A PARTIR DE DADOS DE DISPOSITIVOS VESTÍVEIS.

TCC apresentado ao Curso de Bacharelado em Sistemas de Informação da Universidade Federal de Pernambuco, Centro de Informática, como requisito para a obtenção do título de bacharel em Sistemas de Informação.

Aprovado em: 24/08/2023.

BANCA EXAMINADORA

Profaº. Maíra Araújo de Santana (Orientadora)
Universidade Federal de Pernambuco

Profº. Sérgio Ricardo de Melo Queiroz (Examinador Interno)
Universidade Federal de Pernambuco

Profaº. Maria Karoline da Silva Andrade (Examinadora Externa)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Aos meus amados pais Samuel e Lucicleide, pelo esforço que fizeram para me educar, pelo estímulo na minha carreira acadêmica, pelo enorme apoio e conselhos em diversos momentos, principalmente quando decidi iniciar minha segunda graduação e realizá-la em paralelo a um doutorado.

À minha querida esposa Sophia, pelo carinho, amor e companheirismo no dia a dia. Por me ajudar a ser mais forte na vida e no trabalho, e pela compreensão quando necessitei me dedicar integralmente aos estudos da graduação e do doutorado, onde passava, por um bom tempo, mais de 14 horas consecutivas na universidade.

À minha orientadora Maíra Araújo de Santana, pelas discussões científicas enriquecedoras sobre o tema pesquisa.

Ao Centro de Informática, por ter me dado à oportunidade de estudar em um centro de excelência em informática e com uma ótima estrutura para aprendizagem.

À banca examinadora, por ter aceitado o convite e pela contribuição oferecida a este trabalho.

A todos os pesquisadores de área de aprendizagem de máquina, espalhados por todos os continentes do mundo, que permitiram o desenvolvimento de algoritmos e métodos que foram utilizados neste trabalho.

“O que todos devemos fazer é nos certificar que estamos usando a inteligência artificial de uma maneira que beneficie a humanidade, e não que a deteriore.”

(Timothy Donald Cook)

RESUMO

O Diabetes é uma doença que atinge 537 milhões de pessoas no mundo. A insuficiência de insulina na doença gera um aumento da glicemia, causando enormes danos ao organismo. Devido à grande disseminação dessa doença e de seu estágio antecessor, o pré-diabetes, muito se tem investido em dispositivos que permitam o monitoramento da doença. Atualmente, estudos estão focando na criação de sistemas inteligentes capazes de utilizar dados de dispositivos vestíveis para predição da glicemia. Um passo importante é prever quando pessoas saudáveis se tornam pré-diabéticas, de forma a alertar sobre a presença dessa condição silenciosa e assim tentar evitar a evolução do indivíduo para a Diabetes Melito tipo II. Para alcançar esse objetivo, torna-se imprescindível a descoberta de sinais digitais que possam desempenhar o papel de biomarcadores para a doença, uma área em constante desenvolvimento na atualidade. Nesse sentido, o propósito deste estudo foi avaliar se os sinais digitais obtidos por meio de dispositivos de pulso inteligentes são capazes de estabelecer uma distinção satisfatória entre os padrões de sinal de um indivíduo saudável e de um indivíduo pré-diabético. Para isso foram utilizados dados públicos de leituras da pulseira Empatica E4, coletados de 16 indivíduos. Os dados foram utilizados de duas formas: I) dados brutos processados para janela temporal de um minuto, II) dados que passaram pelo processo de engenharia de atributos. Para tarefa de predição de presença ou ausência da doença foram utilizados algoritmos de aprendizagem de máquina da classe das árvores de decisão, máquinas de vetores de suporte, redes neurais e algoritmos *Ensemble*. Os resultados indicaram que é possível utilizar os dados da pulseira para predição da presença da doença, mas que o uso das leituras brutas dos sensores para o treinamento de modelos gera predições com uma menor acurácia em relação a utilização de atributos que foram filtrados e decompostos. O algoritmo de Random Forest apresentou a melhor acurácia entre os algoritmos testados, 77,4% e 77,9% nos grupos de treino e teste respectivamente, considerando a base com dados brutos; e acurácia 99,07% e 99,30% de para treino e teste na base submetida ao processo de engenharia de atributos. A aplicação da técnica de PCA permitiu a redução do número de atributos em 73% mantendo-se 95% da variabilidade presente na base, isso reduziu o custo computacional mantendo uma acurácia de 95,3% e 95,7% para treino e teste respectivamente. O treinamento deste modelo permitirá criar soluções usando dados de dispositivos vestíveis (pulseiras e relógios inteligentes) para gerar durante a transição para a pré-diabetes, facilitando o estabelecimento de medidas preventivas contra o Diabetes.

Palavras-chave: Inteligência Artificial, Diabetes Mellitus, Estado Pré-Diabético, Dispositivos Eletrônicos Vestíveis.

ABSTRACT

Diabetes is a disease that affects 537 million people worldwide. The deficiency of insulin in this disease leads to an increase in blood glucose levels, causing enormous damage to the body. Due to the great dissemination of this illness and its predecessor stage, pre-diabetes, considerable investment has been dedicated to the development of devices that enable disease monitoring. Currently, studies are focusing on the creation of intelligent systems capable of using data from wearable devices to predict blood glucose. A crucial step is to predict when healthy people become pre-diabetic, to raise awareness of the presence of this silent condition and thus try to prevent the individual from developing into Type II Diabetes Mellitus. To achieve this goal, it is essential to discover digital signals that can play the role of biomarkers for the disease, an area in constant development today. In this sense, the purpose of this study was to evaluate whether the digital signals obtained through smart wrist devices can establish a satisfactory distinction between the signal patterns of a healthy individual and a pre-diabetic individual. For this, public data from sensors of the Empatica E4 bracelet, collected from sixteen individuals, were used. The data were used in two ways: I) raw data processed for a time window of one minute, II) data that went through the feature engineering process. For the task of predicting the presence or absence of the disease, machine learning algorithms from the decision tree class, support vector machines, neural networks and Ensemble algorithms were used. The results showed that it is possible to use wristband data to predict the presence of the disease, but the use of raw sensor readings for model training generates predictions with less accuracy compared to the model training on attributes that were filtered and decomposed. The Random Forest algorithm showed the best accuracy among the tested algorithms, 77.4% and 77.9% in the training and test groups respectively, considering the database with raw data; and accuracies 99.07% and 99.30% for training and testing on the database that went through the feature engineering process. The application of the PCA technique allowed the reduction of the number of attributes by 73%, keeping 95% of the variability present in the base, this reduced the computational cost keeping an accuracy of 95.3% and 95.7% for training and testing, respectively. The training of this model will enable the creation of solutions that use data from wearable devices (bracelets and smart watches) to generate alerts when an individual progresses towards prediabetes, easing the implementation of preventive measures against diabetes.

Keywords: Artificial Intelligence, Diabetes Mellitus, Prediabetic State, Wearable Electronic Devices

LISTA DE ILUSTRAÇÕES

Figura 1 - Dispersão do diabetes melito pelo mundo	22
Figura 2 - Fluxograma para o diagnóstico diferencial do diabetes.....	25
Figura 3 – Modelos de glicosímetros disponíveis comercialmente.....	27
Figura 4 – Tipos de tecnologias não invasivas para a detecção da glicemia.....	28
Figura 5 – Sensores não invasivos utilizados como fonte de dados em estudos de predição de glicemia.	40
Figura 6 – Resumo geral do fluxo de trabalho utilizado neste estudo.....	52
Figura 7 – Histograma dos atributos antes e depois da imputação de dados.	54
Figura 8 – Perfil de correlação e distribuição dos atributos da base de dados.	55
Figura 9 – Resultado da otimização dos hiperparâmetros dos modelos testados.....	57
Figura 10 – Resultado das métricas de avaliação de desempenho dos modelos.	58
Figura 11 – Curvas ROC e DET para cada um dos modelos testados.	59
Figura 12 – Acurácia da predição no conjunto de teste.....	60
Figura 13 – Matrizes de confusão resultantes da predição do conjunto de teste.....	61
Figura 14 - Correlação e distribuição dos atributos das bases após a aplicação do FLIRT e PCA.	63
Figura 15 - Resultado da otimização dos hiperparâmetros do Random Forest nas bases de dados após a aplicação do FLIRT e PCA.....	64
Figura 16 - Resultado da validação cruzada com o algoritmo de Random Forest nas novas bases de dados.	65
Figura 17 – Curvas ROC e DET construída a partir do uso de Random Forest em bases com diferentes dimensionalidades.....	66
Quadro 1 – Principais características de trabalhos relacionados.....	43
Quadro 2 – Descrição dos atributos da base de dados.....	48
Quadro 3 – Hiperparâmetros avaliados para cada um dos algoritmos de aprendizagem de máquina.	51

LISTA DE ABREVIASÕES

AUROC	<i>Area under the receiver operating characteristic</i>
BVP	<i>Blood Volume Pulse</i>
DM1	Diabetes melito tipo I
DM2	Diabetes melito tipo II
DMID	Diabetes melito dependente de insulina
DMNID	Diabetes melito não dependente de insulina
EDA	<i>Electrodermal Activity</i>
HbA1C	Hemoglobina Glicada
HR	<i>Heart Rate</i>
IBI	<i>Interbeat Interval</i>
IoMT	<i>Internet of Medical Things</i>
LOOCV	<i>Leave-one-person-out cross validation</i>
MAD	Desvio Absoluto Mediano
MAE	Erro Absoluto Médio
MAPE	Erro Percentual Médio Absoluto
mARD	Diferença relativa média absoluta
MFCC	<i>Mel frequency cepstral coefficients</i>
ML	<i>Machine Learning</i>
PCA	Análise de Componentes Principais
RMSE	Raiz quadrada do erro-médio
RNP	Rede Neural Profunda

SUMÁRIO

1	INTRODUÇÃO	12
1.1	DELIMITAÇÃO DO TEMA	12
1.2	HIPÓTESE	13
1.3	JUSTIFICATIVA	13
2	REVISÃO DE LITERATURA.....	15
2.1	A GLICEMIA E A INSULINA	15
2.2	O DIABETES	16
2.2.1	<i>O diabetes insípido (diabetes insipidus)</i>	17
2.2.2	<i>O diabetes melito (diabetes mellitus)</i>	18
2.2.3	<i>O pré-diabetes.....</i>	20
2.3	EPIDEMIOLOGIA DO DIABETES E PRÉ-DIABETES	21
2.4	MÉTODOS DE DIAGNÓTICO DO DIABETES MELITO	23
2.5	TECNOLOGIAS UTILIZADAS PARA O ACOMPANHAMENTO DO DIABETES MELITO	24
2.5.1	<i>Métodos invasivos</i>	25
2.5.2	<i>Métodos não invasivos ou minimamente invasivos</i>	28
2.6	INTELIGÊNCIA ARTIFICIAL E MANEJO DA DOENÇA.....	32
2.7	DESENVOLVIMENTO DE BIOMARCADORES DIGITAIS	35
2.8	TRABALHOS RECENTES RELACIONADOS	38
3	OBJETIVOS	46
3.1	OBJETIVO GERAL.....	46
3.2	OBJETIVOS ESPECÍFICOS	46
4	METODOLOGIA.....	47
4.1	FONTE DE DADOS	47
4.2	DESCRIÇÃO DO BANCO DE DADOS.....	47
4.3	PRÉ-PROCESSAMENTO DOS DADOS	47

4.3.1	<i>Pré-processamento para construção de uma base de dados de sinais brutos.....</i>	47
4.3.2	<i>Pré-processamento para construção de uma base de dados de sinais decompostos</i>	
	49	
4.4	ALGORITMOS AVALIADOS.....	50
4.5	VALIDAÇÃO ESTATÍSTICA	51
5	RESULTADOS E DISCUSSÃO	53
5.1	ANÁLISE EXPLORATÓRIA DA BASE CONTENDO OS DADOS BRUTOS DERIVADOS DOS SENSORES.....	53
5.2	ANÁLISE DA PROVA CONCEITUAL E DA QUALIDADE DOS MODELOS PARA CLASSIFICAÇÃO DA PRÉ-DIABETES	58
5.3	ANÁLISE DA BASE CONTENDO OS DADOS DECOMPOSTOS PELO PACOTE FLIRT E APLICAÇÃO DO ALGORITMO DE RANDOM FOREST	62
6	CONCLUSÃO.....	70
	REFERÊNCIAS	71

1 INTRODUÇÃO

1.1 DELIMITAÇÃO DO TEMA

O Diabetes representa um grupo de doenças que acometem a população humana. Dentre elas, a mais comum é a diabetes mellitus tipo II (DM2). Essa doença, que corresponde a cerca de 90% dos casos de diabetes, geralmente se desenvolve lentamente, na maior parte das vezes em adultos e é precedida por uma condição conhecida como pré-diabetes. A DM2 é caracterizada por uma resistência aos efeitos da insulina, um hormônio que responde as concentrações de glicose no sangue. Já no pré-diabetes é possível sentir os efeitos dessa resistência à insulina, o que faz com que a glicemia se eleve e diversos sintomas comecem a surgir (BERG *et al.*, 2019; RODACKI *et al.*, 2022; SOUZA *et al.*, 2012).

O termo glicemia diz respeito à concentração do carboidrato glicose no sangue de indivíduos. A concentração normalmente encontrada num indivíduo em jejum é de 80 a 99 mg de glicose para cada dL de sangue. Como a glicose é a fonte energética básica de um animal, quando cai muito abaixo dessa concentração, as pessoas começam a sentir efeitos adversos, como tonturas e sudoreses, o que é chamado de hipoglicemia, um quadro agudo que é facilmente corrigido com a ingestão de alimentos. Quando a glicemia fica acima do patamar de 99mg/dL de sangue por um período muito grande de dias, começam a surgir complicações crônicas como a pré-diabetes e um tipo específico de diabetes (GUYTON; HALL, 2006).

É recomendado que todo portador do diabetes melito realize aferições diárias de glicose. Por isso, foram desenvolvidos ao longo das últimas décadas, aparelhos denominados glicosímetros. Boa parte dos glicosímetros comercializados é constituída de um aparelho portátil, onde é colocada uma fita de teste contendo uma gota de sangue do paciente. Isso traz um enorme inconveniente para o portador, ele precisa furar a ponta do dedo para extração de sangue a cada vez que precisa mensurar sua glicemia (KHAN *et al.*, 2019). Embora glicosímetros contínuos com microssensores inseridos na pele do paciente estejam cada vez mais se tornando comuns, eles ainda são bastante caros e utilizados apenas por poucos pacientes que possuem a diabetes melito (geralmente portadores de DM1) ou que estão no grupo de risco para seu desenvolvimento, os chamados pré-diabéticos, principalmente em países onde não há políticas de reembolso de parte dos altos custos com esse tipo de glicosímetro (DATYE *et al.*, 2021; SILVA; PATEL, 2023). Isso resulta no fato com que boa parte dos portadores de diabetes não tenha acesso a medição contínua da glicemia. Além disso, os pré-diabéticos também ficam

desalentados pois não há recomendação do uso desses aparelhos para esse grupo de indivíduos (DUAN; KENGNE; ECHOUFFO-TCHEUGUI, 2021).

Visando tornar a medicação da glicemia algo mais fácil e barato, métodos de detecção não-invasivos vêm sendo desenvolvido nas últimas duas décadas. Estes métodos, no entanto, sofrem com uma série de desafios, como conseguir uma alta sensibilidade e especificidade, aplicação para todas as variações anatômicas e fisiológicas individuais. Além de ter um *design* simplista e portátil, com uma boa usabilidade e experiência para o usuário (LIN, 2017). Recentemente, pesquisas vêm sendo realizadas a fim de verificar se dispositivos vestíveis comuns, como pulseiras e relógios inteligentes, têm a capacidade de oferecer informações esclarecedoras sobre o perfil da glicemia dos utilizadores. Caso isso possa ser colocado em prática com boa sensibilidade e especificidade, o número de pessoas com acesso a informações sobre sua curva glicêmica diária expandirá incrivelmente (BENT *et al.*, 2021a). Uma questão a ser resolvida é se os dados fisiológicos capturados por pulseiras inteligentes são suficientes para diferenciação de indivíduos saudáveis e indivíduos que fazem parte do grupo de risco para desenvolvimento da diabetes, os pré-diabéticos, e qual a melhor arquitetura de aprendizagem de máquina para realizar essa previsão.

1.2 HIPÓTESE

Alguns algoritmos de aprendizagem de máquina devem ser capazes de diferenciar indivíduos saudáveis e pré-diabéticos apenas com base no sexo dos indivíduos e em dados fisiológicos medidos de maneira não invasiva.

1.3 JUSTIFICATIVA

Devido ao fato de o diabetes atingir mais de 500 milhões de pessoas atualmente, com um número ainda maior de pessoas portadoras da pré-diabetes com grandes chances de evoluir em para o diabetes (INTERNATIONAL DIABETES FEDERATION, 2021), é necessário o estudo de métodos que diagnostiquem a pré-diabetes precocemente. Uma vez que o tratamento inicial, seja medicamentoso ou de mudança de estilo de vida, pode interromper a progressão para a doença (GIACAGLIA *et al.*, 2022).

Apesar da sua importância, nenhum dos métodos comercialmente disponíveis são focados no público que ainda não é portador da diabetes (LIN, 2017). Apenas nos últimos anos, estudos

vêm buscando encontrar padrões utilizando dispositivos não-invasivos que monitorem com eficiência e que possam expandir o número de pessoas que tem acesso a um monitoramento da glicemia em tempo real (BENT *et al.*, 2021a). Dessa forma faz-se necessário buscar algoritmos que consigam detectar a transição de indivíduos saudáveis para o grupo de pré-diabéticos a partir de dispositivos não invasivos simples e com uma elevada sensibilidade e especificidade.

2 REVISÃO DE LITERATURA

Esta revisão está dividida em 8 partes que abordam desde os conceitos básicos de regulação da glicose, falando sobre o diabetes, passando por métodos tradicionais de diagnóstico e acompanhamento, chegando até novas tecnologias baseadas em sistemas computacionais que auxiliam na prevenção e acompanhamento da doença.

2.1 A GLICEMIA E A INSULINA

A glicemia é um termo utilizado para descrever a concentração de glicose no sangue. A glicose é o principal substrato das reações bioquímicas celulares que liberam energia para os seres vivos (RODWELL *et al.*, 2017). No ser humano há padrões normais de glicemia. Quando uma pessoa está em jejum é adequando que a sua glicemia esteja entre 80 e 90, ou seja, que haja entre 80 e 90 mg de glicose para cada 100 ml de sangue. Após a alimentação esse padrão se altera, mas há alguns limites que não devem ser ultrapassados para que o indivíduo viva com saúde. Esses padrões de concentração de glicose são mantidos graças a Insulina (GUYTON; HALL, 2006).

A insulina é uma proteína de baixo peso molecular, cerca de 5,8 Kd, produzida pelas células β das ilhotas do pâncreas e que atua como um hormônio, estando principalmente associada ao metabolismo de carboidratos num animal, embora também esteja envolvida no metabolismo de proteínas e de lipídeos. Ela é liberada em grandes quantidades no sangue quando uma pessoa realiza a ingestão de uma grande quantidade de alimento, ou de alimentos muito calóricos (sua concentração no plasma sanguíneo aumenta 10 vezes em até 5 minutos após a ingestão). Quando está na corrente sanguínea, ela possui dois destinos, ou é rapidamente degradada pela enzima *insulinase*, ou se liga aos receptores presentes na membrana plasmática das células (GUYTON; HALL, 2006).

A ligação da insulina no receptor na superfície da célula dispara uma sinalização intracelular ao ativar uma proteína tirosina quinase. Os efeitos da cascata de sinalização são diversos e entre eles está, principalmente, o aumento da captação de glicose pelas membranas das células. Isso faz com que a glicose que está em abundância no sangue de vasos e capilares penetre nas células para ser metabolizada, seja para produção imediata de energia como para armazenamento em forma de outras substâncias, como por exemplo o glicogênio (um carboidrato). Isso causa uma redução da glicemia do indivíduo. É justamente sob a forma de glicogênio, que cerca de 60% da glicose obtida logo após as refeições é armazenada no fígado,

sob efeito da insulina. A ação da insulina, além de reduzir os níveis de glicose circulante no sangue, também inibe a gliconeogênese (síntese de glicose a partir de substratos que não são carboidratos) no fígado, impedindo que o corpo produza glicose, uma vez que já há alta disponibilidade de glicose no corpo. Como a insulina participa do metabolismo de diversos compostos, e não apenas da glicose, sua falta não causa apenas problemas de elevação do nível de açúcar no sangue, mas também provoca a conversão de ácidos graxos em colesterol e lipídeos, que são lançados na corrente sanguínea, causando problemas comuns em diabéticos, como a aterosclerose (GUYTON; HALL, 2006). Em resumo a presença da insulina causa principalmente:

- a) Absorção da glicose pelas células e consequentemente a redução da glicemia;
- b) Interrupção dos mecanismos bioquímicos que sintetizam glicose a partir de outros compostos;
- c) Aumento da síntese e armazenamento de proteínas;
- d) Aumento da síntese e armazenamento de gorduras.

Uma vez que há uma estreita relação entre a glicemia e a concentração de insulina, quanto mais glicose disponível no sangue, maior a concentração de insulina, numa relação não linear, mas diretamente proporcional. A concentração de insulina chega ao máximo quando a concentração de glicose fica entre 400 e 600 mg/100 ml de sangue. A baixa da glicemia, como resultado da ação da insulina, estimula a redução da concentração de insulina no sangue. Essa relação contribui para a homeostase corpórea, evitando os efeitos indesejados da alta concentração de glicose, como: I) desidratação celular, II) perda de glicose pela urina, III) aumento da diurese, IV) lesões em tecidos e vasos sanguíneos. Quando essa homeostase é perdida pode haver o surgimento do diabetes melito (GUYTON; HALL, 2006).

2.2 O DIABETES

O diabetes é um conjunto de doenças bastante conhecidas por afetar o ser humano. O termo “diabetes” é muito antigo, assim como a doença. Segundo Areteu, um médico turco que viveu no segundo século da era atual, a doença era vista como “uma desintegração da carne e dos membros em urina”, devido ao fato de os pacientes com essa doença produzirem urina numa quantidade maior que a convencional (BERG *et al.*, 2019). Há dois tipos de doenças que são coletivamente chamadas de diabetes, mas que não estão relacionadas entre si, o diabetes *mellitus* e o diabetes *insipidus*. O primeiro envolve o metabolismo da glicose e a relação com a

insulina, sendo tratado com mais detalhes nos próximos tópicos dessa revisão, já o segundo é uma doença associada ao controle osmótico de fluidos no corpo humano (GUYTON; HALL, 2006). Neste trabalho, para fins de simplificação, o diabetes melito será chamado, a partir desse ponto, apenas de diabetes, enquanto qualquer menção a outra forma de diabetes será feita com o termo diabetes insípido.

2.2.1 O diabetes insípido (*diabetes insipidus*)

O diabetes insípido é uma doença caracterizado por uma alta produção e eliminação de urina por um indivíduo, podendo chegar a níveis maiores que 15 litros diários. Essa urina é geralmente transparente, diferenciando-se da urina convencional, que possui um aspecto amarelado. Além desse sintoma, pacientes com essa doença apresentam uma sede extrema e constante, como um impulso de compensação da perda de água. A exposição deste portador a médios períodos sem água já é capaz de levar a morte por desidratação (LEVY; PRENTICE; WASS, 2019). O diabetes insípido, portanto, não tem nenhuma relação direta com a glicemia do indivíduo ou seus níveis de insulina.

A causa da diabetes insípido está relacionada a uma deficiência na produção ou sensibilidade do hormônio vasopressina, também conhecida como hormônio antidiurético, na neuro-hipófise ou nos rins, respectivamente. A doença pode ser desenvolvida após o desenvolvimento de outras patologias na hipófise, órgão que produz o hormônio. Sem quantidade suficiente desse hormônio, os rins não recebem a sinalização necessária para reabsorver a urina, e passa a excretá-la em excesso, nesse caso a doença é chamada de diabetes insípido central. Outro fator para o desenvolvimento da doença é o consumo excessivo de medicamentos que afetam os rins, ou o surgimento de doenças que afetam a medula renal, nestes casos, o rim danificado não responde mais à vasopressina, e a doença se torna conhecida como diabetes insípido nefrogênico. Menos comumente, causas genéticas e anomalias congênitas, também podem levar ao desenvolvimento dessa doença (GUYTON; HALL, 2006).

A doença é rara, atingindo cerca de 0,004% da população. Seu diagnóstico pode ser feito com base no volume de urina produzida em cada 24 h e pela osmolaridade urinária. Uma outra opção de diagnóstico é feita a partir da dosagem de um peptídeo glicosilado chamado copeptina. Alto níveis desse peptídeo, após a administração de uma solução salina ao paciente, confirmam o diagnóstico de diabetes insipidus, tornando-o um bom biomarcador da doença (REFARDT; WINZELER; CHRIST-CRAIN, 2020).

2.2.2 O diabetes melito (*diabetes mellitus*)

O diabetes melito é uma doença crônica que envolve principalmente o desenvolvimento de níveis elevados de glicose no sangue do indivíduo, estando diretamente ligado a insulina. O termo melito deriva do termo em latim *mellitus*, que significa “adoçado com mel” ou “doce como o mel”, fazendo referência a um dos sintomas da doença, a presença de açúcar na urina dos portadores da doença (BERG *et al.*, 2019). A doença surge quando o pâncreas não é mais capaz de produzir o hormônio, ou quando o corpo não é mais capaz de utilizar a insulina produzida de forma eficaz. O aumento da concentração da glicose leva, além dos efeitos já citados, ao mal funcionamento dos vasos sanguíneos com extravasamento de fluidos para os tecidos, ou problemas na difusão de substâncias, podendo causar infarto, derrame, retinopatia e consequentemente cegueira, isquemia e gangrena. É desejável que a glicemia esteja entre 80 e 90 nos seres humanos. Quando a diabetes é instalada, essa glicemia fica muito mais alta, acima de 126 de acordo com as convenções. Apesar disso, um valor acima de 110 já é um forte alerta de risco para a doença (GUYTON; HALL, 2006).

Além da elevação da glicemia, dada a redução da absorção da glicose pelas células, o indivíduo diabético também passa a ter problemas com o metabolismo de lipídeos e proteínas, já que ambos são diretamente afetados pela insulina, como visto anteriormente (RODWELL *et al.*, 2017).

Como citado, o diabetes melito tem duas causas diferentes, resultando em dois principais subtipos distintos da doença: o diabetes tipo I, que pode também ser conhecido como “diabetes de início juvenil” ou “diabetes melito dependente de insulina” (DMID), ocorre quando a síntese de insulina fica comprometida, já o diabetes tipo II é conhecido como “diabetes de início tardio” ou “diabetes melito não dependente de insulina” (DMNID) é causado por uma redução da sensibilidade das células ao efeito da insulina (GUYTON; HALL, 2006; RODWELL *et al.*, 2017). Além desses dois grandes subtipos do diabetes, há também o diabetes gestacional, além de uma série de subtipos menores. O diagnóstico dos principais subtipos da doença é realizado de forma clínica, pois os indivíduos desenvolvem sintomas muitas vezes distintos (RODACKI *et al.*, 2022).

2.2.2.1 Diabetes tipo I

O diabetes tipo I (DM1) é derivado da destruição das células β produtoras de insulina (no pâncreas), seja por uma doença autoimune, seja por uma outra infecção, ou até mesmo por

herança genética que favorece a destruição dessas células. Em geral essa doença surge até o início da idade adulta. Os indivíduos portadores, por não produzirem insulina, necessitam que o hormônio seja obtido de outras fontes para que eles possam sobreviver, um tratamento chamado de insulinoterapia (BERG *et al.*, 2019; RODACKI *et al.*, 2022).

A falta da insulina nesta forma da doença, além de causar acúmulo de glicose derivada da alimentação no sangue, que não é captada de forma eficiente pelas células, também estimula os tecidos a produzirem mais glicose. Com isso, esses pacientes podem ter uma glicemia aumentada para 300 mg/dL até por volta de 1200 mg/dL, muito maior que a glicemia em uma pessoa saudável. Apesar disso, como não há insulina, as células entendem que devem dar prioridade ao metabolismo de lipídeos, com liberação de substâncias tóxicas no sangue, levando a quadros de acidose metabólica, que rapidamente leva ao coma. Essas alterações metabólicas causadas pela doença obrigam o organismo a ajustar suas respostas fisiológicas, levando a sinais detectáveis alterados, como alterações na frequência da respiração. Os indivíduos com acidose, por exemplo, passam a ter uma respiração rápida e profunda (GUYTON; HALL, 2006).

2.2.2.2 Diabetes tipo II

O tipo II do diabetes é o subtipo mais comum da doença. Cerca de 90% de todos os casos de diabetes no mundo se encaixam nessa categoria. Ela é uma doença que está muito associada ao sobrepeso e a obesidade, se tornando mais comum com o envelhecimento do indivíduo. Mas também há outras causas como, por exemplo, a tendência genética ou a obtenção de síndromes que prejudicam a sinalização da captação da insulina. Nessa doença, diferentemente da DM1, os portadores possuem níveis normais de insulina ou até mesmo níveis elevados, porém as células não respondem mais, criando um fenótipo de resistência à insulina (BERG *et al.*, 2019; GUYTON; HALL, 2006; RODACKI *et al.*, 2022). No entanto, também é possível que alguns indivíduos portadores de DM2 também desenvolvam tardeamente uma deficiência na produção da insulina (RODACKI *et al.*, 2022).

A DM2 não é uma doença que surge de repente, primeiro o paciente entra em um estado conhecido como **síndrome metabólica**, onde a obesidade é o principal fator. Quando a gordura começa a se acumular no fígado e nos músculos, o corpo começa a precisar de cada vez mais insulina. Dessa forma, nesse período desenvolve-se a resistência à insulina, além de uma

elevação dos padrões normais de glicemia, o desenvolvimento da dislipidemia¹, e o surgimento da hipertensão. Essa resistência à insulina faz com que cada vez mais a célula β precise produzir o hormônio, mas há um limite para isso. Chega um momento em que a quantidade de insulina produzida se acumula em grandes quantidades no retículo da célula, causando uma toxicidade. O corpo não consegue mais atuar para baixar a glicemia e o paciente desenvolve a diabetes, necessitando do uso de drogas que agem reduzindo a resistência à insulina, não havendo recuperação para o paciente. Quando a célula não consegue mais lidar com a grande quantidade de insulina, então ela entra em apoptose (morte celular). A partir desse momento o DM2 está totalmente desenvolvido e avançado, apenas o tratamento para controle da doença pode ser realizado, com os medicamentos e com a administração de insulina (BERG *et al.*, 2019; GUYTON; HALL, 2006). Com isso, fica claro que o tipo 2 da diabetes surge de uma forma bastante distinta da DM1, sendo uma progressão lenta onde há um estágio anterior chamado de pré-diabetes.

2.2.3 O pré-diabetes

O estágio do pré-diabetes é uma fase em que as pessoas ainda não possuem os níveis de glicemia alterados o suficiente para possuírem a doença. Esses níveis, contudo, não são normais, ou seja, não deveriam ocorrer em pessoas saudáveis. No pré-diabetes estão inseridas as pessoas que possuem a glicemia de jejum alterada (100 – 125 mg/dL) e uma sensibilidade reduzida a glicose (SOUZA *et al.*, 2012). Apesar disso, os valores de corte dos testes para a condição são diferentes entre diferentes associações de saúde. Eles divergem por exemplo entre a Organização Mundial da Saúde e a Associação Americana de Diabetes. Dessa forma a definição e tamanho do grupo de pessoas pré-diabéticas varia de acordo com o autor de um estudo e o instituto que coordena algum levantamento. No entanto, o grupo de pré-diabéticos incluem pessoas que possuem certo grau de obesidade e resistência à insulina, assim como pessoas com algum grau de deficiência das células β do pâncreas (BEULENS *et al.*, 2019).

De acordo com a Associação Americana de Diabetes, o pré-diabetes não deve ser considerado uma entidade clínica, mas uma faixa de risco para o diabetes e a doença cardiovascular, uma vez que a condição está frequentemente associada a obesidade (AMERICAN DIABETES ASSOCIATION, 2021). Indivíduos que desenvolvem o pré-

¹ Elevação das concentrações sanguíneas de triglicerídeos, colesterol e lipoproteínas de baixa densidade.

diabetes possuem alto risco de evoluir para o diabetes. No geral, apenas 25% das pessoas com essa condição conseguem se recuperar do problema e afastar a doença. Esse risco aumenta com a idade do indivíduo e é influenciado pela sua glicemia em jejum. Indivíduos mais idosos possuem ainda menos chance de se recuperar do pré-diabetes, assim como aqueles cuja glicemia se encontra entre 110 e 125 (limite para diabetes). A síndrome metabólica desenvolvida no pré-diabetes, também é um fator crucial para o ponto de não retorno, onde a doença está estabelecida (GIACAGLIA *et al.*, 2022).

Para o tratamento do pré-diabetes é recomendado que os indivíduos mudem seu estilo de vida, pratiquem mais exercícios físicos, reduzam o consumo de alimentos calóricos e consequentemente seu peso. Apesar disso, em alguns casos o tratamento farmacológico é necessário para que a pessoa não evolua para o diabetes (GIACAGLIA *et al.*, 2022). No entanto, o grande problema para o início do tratamento é o diagnóstico da doença, uma vez que a pré-diabetes é uma condição silenciosa.

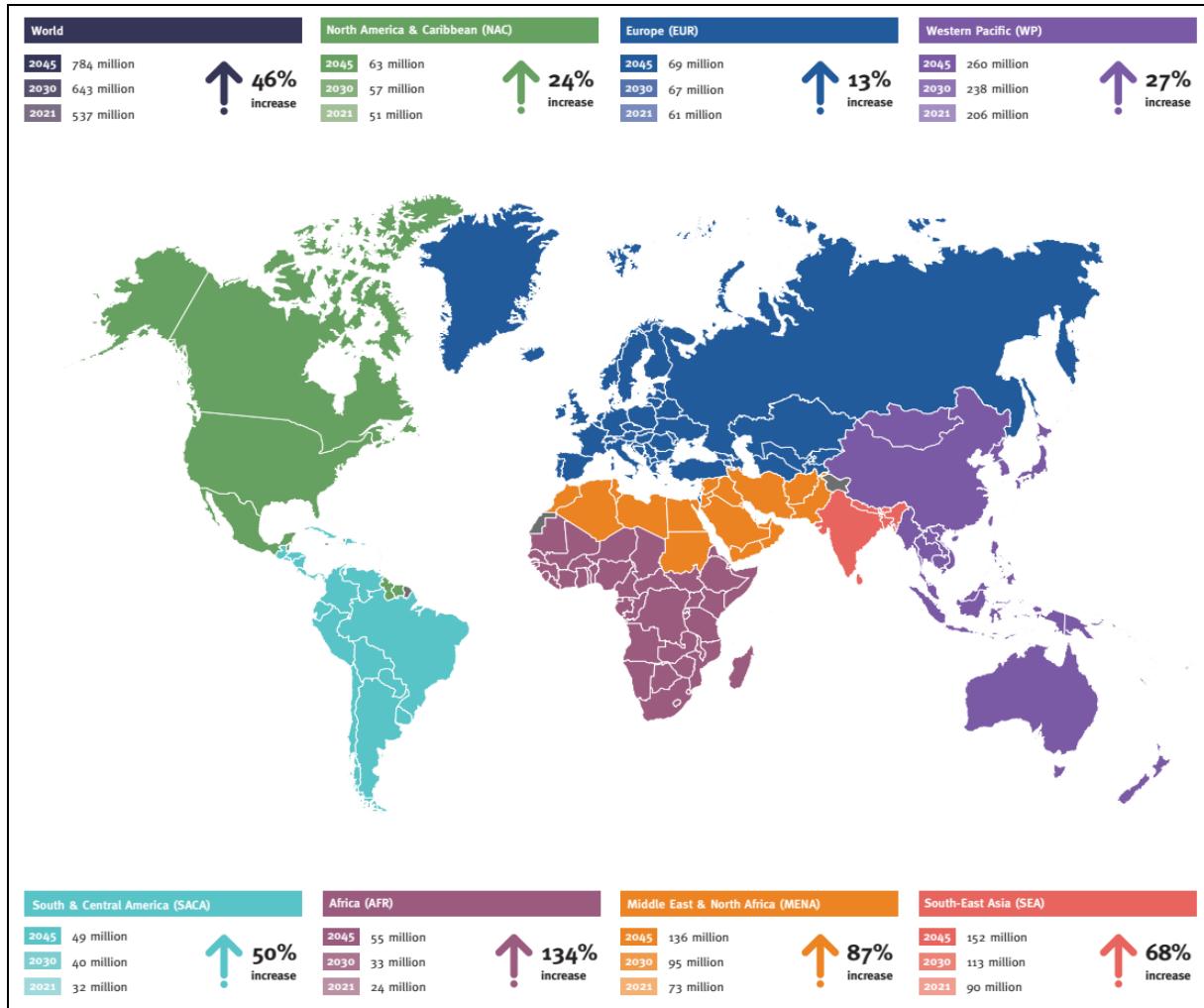
Diversos estudos criaram modelos de risco para predição do pré-diabetes baseado em uma série de características de cada indivíduo. Esses modelos geralmente levam em considerações características que não demandam a realização de testes sanguíneos pelos indivíduos, mas sim o preenchimento de questionários padronizados, que contém características básicas das pessoas e que podem ser boas preditoras para o risco de diabetes. Quando a pontuação final do questionário indica a existência de um alto risco para o desenvolvimento da doença, é então recomendado que o indivíduo passe a realizar exames laboratoriais para acompanhamento médico (BEULENS *et al.*, 2019). Geralmente as características preditoras básicas mais presentes nos questionários são: Idade, índice de massa corpórea (IMC), presença de hipertensão e histórico familiar de diabetes (BARBER *et al.*, 2014).

2.3 EPIDEMIOLOGIA DO DIABETES E PRÉ-DIABETES

O diabetes é uma das grandes doenças do século 21. A Federação Internacional de Diabetes estimou um número de doentes em 537 milhões de pessoas adultas (20-79 anos) em 2021, ou seja, uma em cada dez pessoas no mundo possui a doença (Fig. 1). Em 2045 esse número pode chegar a 784 milhões, o que corresponde a uma em cada 8 habitantes do planeta. Isso vem disparando um alarme para o rápido avanço da doença, que só no ano de 2021 já causou a morte de 6,7 milhões de pessoas. Esse número é impressionante, pois quer dizer que a cada 5 segundos, uma pessoa morre de diabetes na Terra. Outro fator assustador em relação

a doença, é que o número de pessoas portadoras de resistência à insulina, também conhecido como pré-diabetes, no ano de 2021 superou o número de diabéticos. A Federação estima que cerca de 541 milhões de pessoas estavam com pré-diabetes em 2021. O que implica em mais de 1 bilhão de pessoas envolvidas com a doença (INTERNATIONAL DIABETES FEDERATION, 2021).

Figura 1 - Dispersão do diabetes melito pelo mundo.



Fonte: International Diabetes Federation (2021)

Legenda: São mostrados os números de portadores da doença ao longo de 8 regiões mundiais, assim como provável número de portadores para os anos de 2030 e 2045.

A prevalência do diabetes é um pouco maior em homens que em mulheres, embora os valores sejam muito parecidos, 10,8% em homens e 10,2% em mulheres, e cresce conforme a faixa etária analisada aumenta. A doença é mais comum em áreas urbanas que em áreas rurais, e atinge principalmente países com renda média ou baixa, onde vivem 432,7 milhões de doentes. Sua maior prevalência em relação a poluição regional encontra-se no Oriente Médio e norte da África, onde cerca de 16% da população possui diabetes. O ranking dos países com o

maior número de doentes é liderado pela China, com cerca de 141 milhões de doentes. O Brasil, no entanto, não está numa situação favorável, com 15,7 milhões de doentes, ele é o sexto do ranking (SUN *et al.*, 2022).

2.4 MÉTODOS DE DIAGNÓSTICO DO DIABETES MELITO

Para uma pessoa ser diagnosticada com diabetes é preciso que sua glicemia esteja acima de um padrão estabelecido pelos estudiosos do diabetes. Quando não se atinge esse limiar, mas a glicemia é maior que a os parâmetros normais, classifica-se o indivíduo como pré-diabético. Estes padrões existem em diversos tipos de testes, entre eles estão os testes de:

- a) Teste de glicemia em jejum: Um teste realizado após pelo menos 8h de jejum de alimentos ou bebidas que podem alterar o metabolismo da glicose. Neste teste, o padrão inferior para diabetes corresponde a uma glicemia de 126, ou seja, uma pessoa cuja proporção de glicose no plasma sanguíneo esteja acima de 126 mg/dL pode ser considerada diabética. Caso este valor esteja entre 100 e 125, a pessoa pode ser considerada pré-diabética (AMERICAN DIABETES ASSOCIATION, 2021; KHAN *et al.*, 2019).
- b) Teste oral de tolerância à glicose: neste teste a glicemia do paciente é aferida duas horas após a ingestão de 75 g de glicose. É considerado saudável o paciente cuja glicemia esteja abaixo de 140, e diabético o paciente cuja glicemia esteja acima de 200. O intervalo entre 140 e 200 é atribuído ao pré-diabetes (AMERICAN DIABETES ASSOCIATION, 2021). Este é o teste padrão ouro para classificação do diabetes, porém requer 8h de jejum prévio à ingestão das 75g de glicose, além de um acompanhamento de uma espera de duas horas onde o paciente deve permanecer em repouso no laboratório. Sendo dos 3 tipos de exames, aquele com maior desconforto para o paciente (DUAN; KENGNE; ECHOUFFO-TCHEUGUI, 2021).
- c) Hemoglobina Glicada (HbA1C): Neste teste é verificada a proporção da proteína hemoglobina presente nas células sanguíneas que está ligada a glicose. Ele é capaz de identificar uma hiperglicemia crônica, uma vez que não sofre variações diárias como o teste de tolerância à glicose ou o teste de glicemia de jejum. Apesar disso, esse teste é indireto, pois não afere diretamente a quantidade de glicose no plasma sanguíneo. Um valor maior que 6,5% é um indicativo de diabetes, enquanto um valor entre 5,7% e 6,4% indica que o indivíduo está na faixa do pré-diabetes (KHAN *et al.*, 2019).

Nem sempre há uma equivalência entre os três testes, ou seja, é possível que um dos testes classifique uma pessoa como diabética, enquanto outro teste não classifique. Nestes casos os testes devem ser repetidos e deve haver pelo menos dois testes distintos da mesma amostra, ou o mesmo teste, mas em duas amostras distintas, para classificar alguém como portador de diabetes. O teste oral de tolerância a glicose classifica mais pessoas na categoria de diabetes ou pré-diabetes que os outros dois testes. Se o teste da hemoglobina glicada variar em relação aos testes de detecção de glicemia, deve-se considerar esses últimos como portadores de maior acurácia, uma vez que pode haver alguma interferência no teste de aferição de HbA1C e uma vez que a HbA1C é uma forma indireta de aferição da glicemia. Apesar disso o teste de HbA1C é mais prático, pois não requer um jejum anterior de pelo menos 8 horas, ou que o paciente fique pelo menos 2 horas aguardando em um laboratório para realizar um exame após ingerir os 75g de glicose (AMERICAN DIABETES ASSOCIATION, 2021).

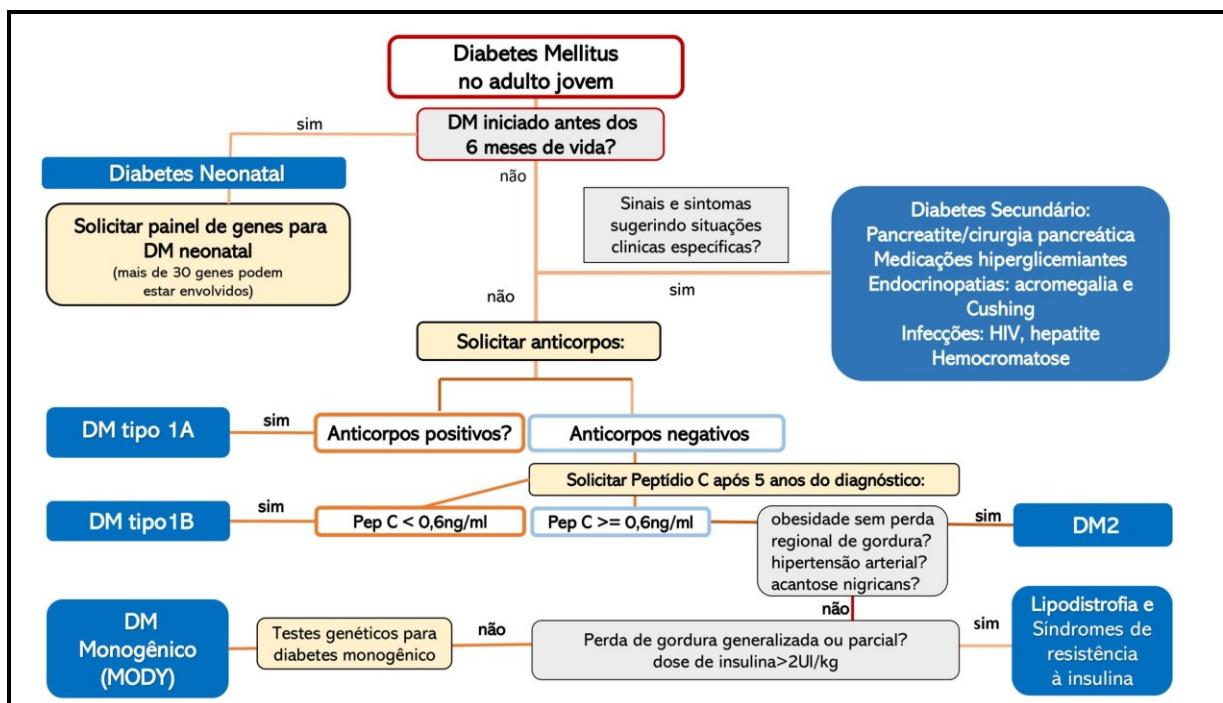
Além de diagnosticar a diabetes, também é importante que o médico saiba diferenciar entre os principais subtipos da doença. O diagnóstico clínico muitas vezes é capaz de diferenciar o diabetes tipo I e o tipo II. Em casos de diagnóstico diferencial, pode-se realizar uma dosagem do peptídeo C para avaliar a função secretória das células β pancreáticas, e também da presença de autoanticorpos que atacam essas células (RODACKI *et al.*, 2022). A Figura 2 mostra um fluxograma estabelecido pela Sociedade Brasileira de Diabetes para o diagnóstico diferencial da doença. Apesar de métodos de diagnóstico diferencial para o diabetes, não há um modelo para o caso do pré-diabetes.

2.5 TECNOLOGIAS UTILIZADAS PARA O ACOMPANHAMENTO DO DIABETES MELITO

De acordo com recomendações de diversos institutos internacionais, os indivíduos portadores de diabetes devem realizar acompanhamento regular do seu nível de glicose, uma vez que a hiperglicemia traz diversos efeitos nocivos. Por outro lado, o uso de medicamentos por parte desses pacientes pode causar momentos de hipoglicemia, que pode levar à morte, o que também comprova a necessidade de acompanhamento constante dos níveis de glicose sanguínea. Uma das recomendações do Instituto Nacional de Saúde dos Estados Unidos é realizar essa verificação diariamente em 4 períodos distintos (KHAN *et al.*, 2019). Para indivíduos que não possuem a doença, mas que já estão no grupo de pré-diabéticos, o teste também é necessário. Segundo a Associação Americana de Diabetes, indivíduos desse grupo devem ser testados pelo menos uma vez ao ano (DUAN; KENGNE; ECHOUFFO-

TCHEUGUI, 2021). Alguns estudiosos, no entanto, argumentam que pelo fato de muitos pré-diabéticos estarem evoluindo para a doença, seria necessário que a verificação dos níveis de glicemia fosse bem mais recorrente (BENT *et al.*, 2021a). Indivíduos não portadores da doença e que também não estão no grupo dos pré-diabéticos também precisam acompanhar o seu nível de glicemia, porém numa frequência bem menor. Não há uma frequência exata para isso, mas simulações mostram que indivíduos com mais de 45 anos devem testar sua glicemia pelo menos uma vez a cada 3 anos (DUAN; KENGNE; ECHOUFFO-TCHEUGUI, 2021), o que pode ser feito com exames básicos laboratoriais sem a necessidade de tecnologias de acompanhamento de glicose.

Figura 2 - Fluxograma para o diagnóstico diferencial do diabetes.



Fonte: Rodacki *et al.* (2022)

2.5.1 Métodos invasivos

Os métodos tradicionais de mensuração diária de glicose para indivíduos diabéticos envolvem o uso de glicosímetros. Estes aparelhos são acompanhados de tiras de teste que são inseridas nos aparelhos logo após o depósito de uma gota de sangue do paciente sobre elas. O resultado é rapidamente mostrado no *display* do aparelho e o paciente pode anotar e guardar essa informação (KHAN *et al.*, 2019). Um exemplo de glicosímetro tradicional é mostrado na Figura 3A.

Um dos inconvenientes dos glicosímetros é a necessidade de furar a extremidade de um dos dedos para realização da extração da gota de sangue toda vez que deve ser feita a aferição da glicemia. Isso causa desconforto e pode gerar cicatrizes no longo prazo, o que pode levar a certa resistência por parte de alguns indivíduos portadores de diabetes (KHAN *et al.*, 2019).

A necessidade de acompanhar a glicemia por um período contínuo e regular, juntamente com a necessidade de evitar perfurações em diversos momentos do dia fomentou o desenvolvimento de dispositivos com sensores implantáveis, que se comunicam com aparelhos externos (decodificadores) sem nenhum uso de cabeamento. Um dos primeiros aparelhos desse tipo capaz de realizar leituras contínuas por pelo menos uma semana e vendido comercialmente foi o Dexcom G4 PLATINUM (Fig. 3 D). Lançado no final de 2013, ele utilizava tecnologia *bluetooth* para realizar a conexão entre o sensor e o decodificador. O sensor implantável era capaz realizar uma leitura a cada 5 minutos, num total de 288 leituras diárias, com uma alta correlação ($R = 0,95$) entre a glicose predita e a glicose real. Sendo também, capaz de emitir alertas de hipo e hiperglicemia. O uso desse aparelho melhorou o controle de glicemia de 83% de um total de 74 indivíduos avaliados num estudo. Apesar disso, o aparelho necessitava de duas calibrações diárias, ou seja, o usuário precisava fornecer diariamente, a cada 12 horas, seu nível de glicose, e para isso deveria utilizar o glicosímetro convencional retirando uma gota de sangue da ponta do dedo, o que causava incômodo (NAKAMURA; BALO, 2015; REDA *et al.*, 2021).

Posteriormente foi lançado comercialmente o aparelho FreeStyle Libre (Fig. 3E), que é vendido até os dias atuais. Seu sensor é capaz de durar o dobro do tempo do Dexcom G4 PLATINUM, contando com uma auto calibração, o que tornou o dispositivo muito mais amigável, fornecendo uma melhor experiência do usuário. Além disso, não há diferença significativa na predição da glicose, entre esses dois dispositivos, na maior parte dos casos de glicemia, exceto quando os níveis de glicose mudavam rapidamente. Neste caso, o resultado do FreeStyle Libre se saiu melhor que o do Dexcom G4 PLATINUM (BOSCARI *et al.*, 2018; REDA *et al.*, 2021). Apesar desses dois aparelhos atingirem uma boa predição de glicose, o seu sensor dura pouco tempo, devendo ser trocado em períodos extremamente curtos, gerando um incômodo para paciente. Por isso, diversas empresas focaram seus esforços para conseguir criar um sensor implantável que durasse por um longo período. Destes, o sensor com maior durabilidade é o Eversense XL (Fig. 3F), que pode durar até 6 meses. Embora o tempo entre as trocas a do sensor, o que exige a ida do paciente a uma clínica com a presença de pelo menos 2 profissionais de saúde, seja maior que o tempo exigido por sistemas anteriores, este monitor

contínuo de glicose continua sendo recomendado apenas para pessoas portadoras de diabetes, visando o controle da hipoglicemia após a administração de medicamentos e a obtenção das metas de tratamento (DEISS *et al.*, 2019).

Figura 3 – Modelos de glicosímetros disponíveis comercialmente.



Fonte: Reda *et al.* (2021)

Legenda: (A) Glicosímetros tradicionais, (B) O primeiro glicosímetro não invasivo e de uso contínuo, chamado GlucoWatch®, (C) um Holter de glicose chamado de iPro2, desenvolvido para acompanhamento médico da glicose, (D) um glicosímetro de medição contínua chamado Dexcon G4, que utilizava um aparelho receptor desenvolvido especificamente para processar e armazenar os dados das leituras de glicose, (E) um glicosímetro de

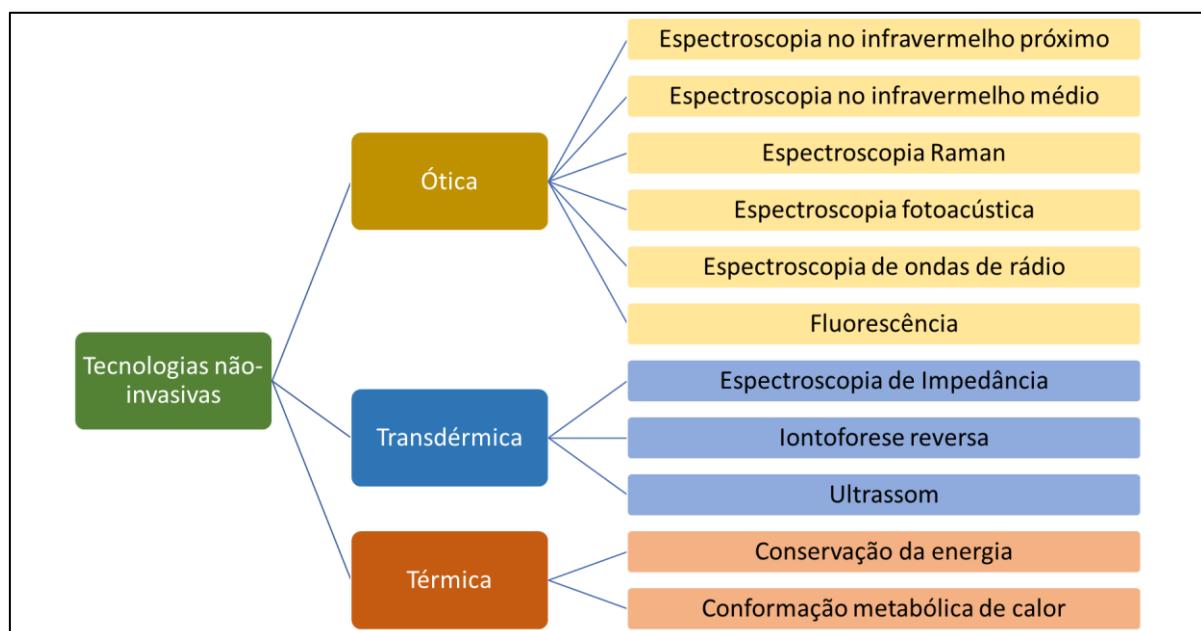
medição contínua chamado FreeStyle, que envia dados para um smartphone, (F) o dispositivo Eversense XL, cujo sensor dura cerca de 6 meses, (G) um glicosímetro de medição contínua, Medtronic 670G, que também é acoplado a um dispositivo de liberação de insulina.

Apesar de o Eversense XL exigir a troca de sensor num prazo bastante longo, uma pequena parte dos usuários relataram algum tipo de irritação na pele após procedimentos de inserção ou retirada do sensor, ou até mesmo uma infecção leve que exigiu o uso de antibióticos (SANCHEZ *et al.*, 2019). Hoje os métodos invasivos de medição contínua estão evoluindo para não causar nenhum tipo de dor ou desconforto ao usuário. Isso está sendo alcançado com o uso de materiais flexíveis, miniaturizados e métodos cada vez mais minimamente invasivos. Sendo cada vez mais fáceis de usar e geralmente transmitindo dados através de redes sem fio para dispositivos de processamento de dados como um aparelho específico ou um smartphone. Além disso, os métodos de medição contínua estão cada vez mais buscando fonte alternativas além da captação de plasma sanguíneo (REDA *et al.*, 2021).

2.5.2 Métodos não invasivos ou minimamente invasivos

Visando eliminar o desconforto causado pelas picadas para retirar gotas de sangue, outras fontes de dados para os glicosímetros passaram a ser estudadas. Geralmente estes dados derivam de tecnologias óticas, transdermais e termais (Fig. 4) (LIN, 2017).

Figura 4 – Tipos de tecnologias não invasivas para a detecção da glicemia.



Fonte: Adaptado de Lin (2017)

Talvez a fonte de dados mais básica para uma coleta não invasiva seja a pele, e por isso, desde o século passado, pesquisadores vêm desenvolvendo métodos que utilizam essas tecnologias para tentar estimar a concentração de glicose a partir deste órgão. O primeiro método não invasivo de medição da glicemia aplicado comercialmente foi utilizando a **tecnologia transdérmica** de iontoporese reversa. O aparelho, desenvolvido na década de 1990, era chamado GlucoWatch®. Ele era colocado no pulso, similarmente a um relógio. Esse aparelho operava através de um método conhecido como iontoporese reversa, onde uma corrente elétrica de baixa intensidade é aplicada gerando o transporte de uma substância através da pele para fora do corpo. No caso do GlucoWatch, esse composto era a glicose, sendo medida por biossensores amperométricos presentes na face do equipamento que ficava em contato com a pele. Todo o ciclo de medição durava cerca de 20 minutos. Os dados de leitura da glicose eram integrados a dados de variação brusca de temperatura e quantidade de transpiração, capturados respetivamente por sensores de temperatura e condutividade da pele (TIERNEY *et al.*, 2000). Um estudo com 92 indivíduos mostrou uma boa correlação ($R = 0,88$, $R^2 = 0,77$) entre a concentração de glicose aferida por glicosímetros tradicionais e através do método de iontoporese (TAMADA *et al.*, 1999). Apesar de ter uma boa correlação com os dados reais, o equipamento não tinha uma correlação extremamente alta. Estudos com um número maior de pacientes diabéticos mostrou um RMSE equivalente a cerca de 43 mg/dL entre os dados reais e o estimado pelo aparelho, representando um desvio moderado entre a glicemia real e a estimada (TIERNEY *et al.*, 2001). Além disso, seus sensores não duravam muito tempo, eles conseguiam realizar leituras por 12 horas. Dessa forma, diversos outros métodos de estimativa da concentração de glicose através da pele foram sendo desenvolvidos nos últimos 20 anos. Apesar disso, o método de iontoporese reversa ainda é utilizado como princípio em novos equipamentos de medição não invasiva, como o SugarBEAT, que permanece em desenvolvimento e utiliza um dispositivo adesivo na pele, que conectado com smartphone, faz leituras a cada intervalo de 5 minutos (YARMAN; KURBANOGLU; SCHELLER, 2020).

Um outro método de tecnologia transdérmica, a espectroscopia de impedância, foi utilizado no aparelho Pendra®, lançado comercialmente no início dos anos 2000, este aparelho saiu de linha poucos anos depois, já que seu uso na população e estudos posteriores revelaram uma baixa correlação entre a glicemia real e a predita ($R=0,35$) (WENTHOLT *et al.*, 2005).

Em relação às **tecnologias térmicas**, um dos métodos de mensuração da glicose é a conformação metabólica de calor, nele a glicemia é deduzida a partir da quantidade de dissipação de calor, taxa de fluxo sanguíneo do tecido local e do grau de saturação de oxigênio

no sangue. Isso é possível, uma vez que o grau de calor dissipado pelo corpo está correlacionado com a quantidade de glicose e oxigênio, já que estes dois estão ligados à geração de energia. Um sistema criado por Tang *et al.* (2008), utilizava três sensores de temperatura, dois sensores de umidade, um sensor infravermelho e um dispositivo de medição óptica para estimar esses parâmetros. Uma função matemática desenvolvida para o sistema então calculava a glicemia baseada no calor, na taxa de fluxo sanguíneo e no grau de saturação de oxigênio no sangue. O coeficiente de correlação entre os dados glicêmicos preditos por esse dispositivo e os dados reais foi de $R=0,856$, um valor muito similar aos métodos anteriores criados nos anos na primeira década do século XXI. O outro método de tecnologias térmicas, a conservação do calor, foi utilizado em alguns equipamentos desenvolvidos nos últimos 10 anos por Zhang *et al.* (2017). Apesar de seu uso juntamente com tecnologias de aprendizagem de máquina, não houve grande aumento no coeficiente de correlação ($R = 0,885$) entre os valores reais e os preditos, em comparação a tecnologias anteriores.

As tecnologias não-invasivas mais utilizadas são com certeza as **tecnologias ópticas**. Embora elas sofram com o baixo sinal produzido pelas moléculas de glicose, o que pode gerar baixa sensibilidade e especificidade, elas são utilizadas por vários aparelhos registrados para comercialização, como: OrSense NBM-200G, C8 MediSensors, Combo Glucometer e Gluco Wise. O uso de luz infravermelha para leitura da glicemia é uma dessas possibilidades. Com esse tipo de onda, podem ser utilizados o infravermelho médio e o infravermelho próximo, em técnicas chamadas de espectroscopia no infravermelho médio e espectroscopia no infravermelho próximo (*near-infrared spectroscopy – NIR*) (LIN, 2017). Nesta última, um comprimento de onda de 900 a 1100 nm é emitido sobre a ponta do dedo, evitando a necessidade da perfuração. Com isso a glicose é medida opticamente, com base na dispersão e absorção de luz através do sangue. Lawand, Parihar e Patil (2015) criaram um sistema onde um LED infravermelho emite uma luz no comprimento de onda de 950 nm sobre a ponta do dedo de um indivíduo. Um fototransistor converte então a luz em tensão. Essa, por sua vez, é utilizada numa regressão polinomial, que prediz a concentração de glicose. Quando comparado com a aferição de glicose via glicosímetro tradicional, esse método apresenta um erro médio de 3% e mostra uma altíssima correlação entre os níveis reais e os preditos de glicose ($R^2 = 0,93$). Apesar disso, o trabalho de Lawand, Parihar e Patil (2015) avaliou apenas níveis de glicose inferiores a 100, um valor mais baixo do que geralmente é encontrado em diabéticos e pré-diabéticos .

Diversos fatores tornam as abordagens não-invasivas muito complexas, como um alto grau de ruído nos dados, derivado do baixo sinal de leitura de glicose; um alto nível de sudorese,

que atrapalha a leitura de alguns sensores; e a interferência do meio externo, por exemplo excesso de frio no ambiente de leitura (LIN, 2017). Dessa forma, além da pele, estudos também focam em fontes alternativas. Uma delas é a secreção lacrimal. No entanto, a quantidade de glicose na lágrima é muito menor que sua concentração no plasma sanguíneo. Um estudo comparando 7 conjuntos de glicosímetros + tiras de teste identificou que apenas 1 desses sistemas poderia ser utilizado como um instrumento de teste para a glicose em indivíduos. Contudo esse estudo validou sua análise apenas em indivíduos saudáveis, não havendo nenhum tipo de teste com secreção lacrimal de pessoas com diabetes ou pré-diabetes (CHA *et al.*, 2014). Além disso, não é simples produzir lágrimas em todo momento que se queira aferir a glicose. Há também estudos que buscam criar sensores a partir de reações enzimáticas que ocorrem com o suor do indivíduo, como o caso de tecidos vestíveis com sensores colorimétricos, que poderiam indicar alterações na glicemia e concentração de ureia (PROMPHET *et al.*, 2020). Apesar da facilidade de obtenção do suor, dispositivos que o utilizam geralmente sofrem com problemas de precisão e estabilidade (REDA *et al.*, 2021).

Com o avanço da tecnologia e o desenvolvimento de múltiplos métodos não invasivos de aferição da glicemia (nem sempre tão eficazes), alguns pesquisadores propuseram unir diversos sensores e métodos de detecção, a fim de reduzir as limitações de especificidade de cada tecnologia (LIN, 2017). Cada método consegue estimar independentemente a glicemia, e a glicemia final pode ser dada pela média ponderada dos resultados de múltiplos sensores. Dessa forma, a estimativa da glicemia baseia-se em vários parâmetros fisiológicos dos pacientes. Apesar de parecer uma simples tarefa de extrair as médias das previsões, deve-se levar em consideração diversos tipos de erro podem interferir na leitura de um ou em parte dos sensores. Dessa forma, em boa parte dos casos, uma simples média dos resultados dos sensores pode não dar o melhor resultado possível. Além disso, uma mudança nos níveis de glicose não reflete instantaneamente nos níveis de diversos parâmetros fisiológicos utilizados. Dessa forma, a estimativa da glicemia deve levar em conta o tempo de atraso entre as mudanças, aumentando em muito a complexidade de todas as análises, de forma que o uso de algoritmos de aprendizagem de máquina vem sendo cada vez mais necessários, para corrigir esse tempo de atraso e dar resultados mais próximos à realidade (GENG *et al.*, 2017).

Uma outra limitação a ser superada pelos dispositivos não invasivos de verificação da glicemia é sua adequação a cada indivíduo específico, uma vez que os parâmetros fisiológicos e anatômicos de cada indivíduo sofrem variações mesmo no seu estado saudável, como por exemplo a cor e espessura da pele. Para superar essa limitação, cada aparelho deve passar por

processos de calibração, onde as leituras reais de glicose devem ser levadas em consideração no cálculo da glicose estimada. Do lado do sistema, isso aumenta bastante a complexidade do algoritmo que será utilizado para compensação. Do lado do usuário, isso pode reduzir a qualidade da experiência de uso, de forma que as tecnologias atuais de medição buscam simplificar bastante a etapa de calibração dos aparelhos e reduzir sua frequência. Isso é feito com sensores e algoritmos cada vez mais robustos, necessitando cada vez mais de inteligência artificial (LIN, 2017).

2.6 INTELIGÊNCIA ARTIFICIAL E MANEJO DA DOENÇA

O diabetes tipo II é uma doença evitável, mas o diagnóstico do problema deve ser realizado cedo, quando a pessoa ainda está na fase do pré-diabetes. Contudo, como visto anteriormente, essa condição é silenciosa. Mapear uma população inteira com exames como teste de glicemia de jejum, teste oral de tolerância à glicose, ou medição da HbA1C é inviável pois isso necessita a retirada e processamento de amostras sanguíneas de milhões de indivíduos, o que seria extremamente custoso e enfrentaria diversas barreiras de aceitação. Um estudo conduzido na Inglaterra estimou que os custos de para diagnóstico da DM2, no ano de 2012, iniciavam em cerca £ 450 (libras esterlinas) para cada caso positivo detectado. Esses custos variavam de acordo com o esquema de diagnóstico realizado. O menor custo foi encontrado apenas quando inicialmente era realizada uma pré-triagem com um método de diagnóstico não invasivo, e só num segundo estágio era realizado algum teste sanguíneo. O esquema de diagnóstico cujo teste sanguíneo era realizado inicialmente chegou a um custo de até 1639 libras por caso positivo, um valor altíssimo (KHUNTI *et al.*, 2012). Isso torna claro a necessidade de existência de algum método que possa fazer uma pré-seleção de um grupo de pessoas para, só então, selecionar dentro desse grupo, aqueles com alterações relacionadas elevação da glicemia ou mesmo resistência à insulina.

Essa abordagem de pré-triagem serve para diagnosticar a possibilidade de pré-diabetes. Para isso, geralmente eram utilizadas as ferramentas de cálculo de risco. Contudo era necessário encontrar quais as melhores variáveis capazes de prever se uma determinada pessoa pode evoluir para a DM2, e para isso diversas ferramentas de predição foram desenvolvidas. Barber *et al.*, (2014) revisaram mais de 250 artigos, e encontraram 18 ferramentas destinadas ao diagnóstico específico do pré-diabetes. A maior parte dessas ferramentas (11) utilizavam um método estatístico clássico para predição, a regressão logística. O restante utilizava algum método de aprendizagem de máquina, como árvores de decisão (6 ferramentas) e máquina de

vetores de suporte (SVM) (1 ferramenta). Não só os métodos de predição se diferenciavam entre as ferramentas, mas também os tipos e a quantidade de variáveis. Quase metade dessas ferramentas não citavam nenhum tratamento de dados faltantes, o que é essencial para dados coletados de pacientes. Além disso, essas ferramentas também se diferenciavam no quesito do responsável pelo seu uso, enquanto algumas poderiam ser utilizadas pelos próprios indivíduos como um autoteste, outras deveriam ser aplicadas apenas por profissional de saúde habilitado. A aplicação de parte dessas ferramentas no mundo real se torna ainda difícil pois 61% delas possuíam falhas no quesito de validação, e um número ainda maior não possuía função de calibração.

Uma dessas ferramentas iniciais que utilizavam aprendizagem de máquina para calcular o risco de diabetes e pré-diabetes era o *Diabetes Risk Calculator*. Ele utiliza a árvore de decisão CART e as variáveis: idade, circunferência da cintura, presença de diabetes gestacional, altura, raça/etnia, hipertensão, história familiar de diabetes e exercício físico, para classificar o indivíduo em 3 níveis de risco: I) diabetes ou pré-diabetes, II) pré-diabetes, III) risco baixo. No entanto seu nível de acerto não era tão grande, com sensibilidade e especificidade de 88% e 75%, respectivamente, para o diagnóstico de diabetes. Mas apenas sensibilidade e especificidade de 75% e 65%, respectivamente, para o diagnóstico de pré-diabetes (HEIKES *et al.*, 2008).

Diversos algoritmos de aprendizagem de máquina podem ser utilizados para classificar uma determinada instância com base em suas características, ou seja, dado um conjunto de características de um indivíduo, poderíamos rotulá-lo como predisposto ou não-predisposto a doença. Isso também poderia ser feito para colocar o indivíduo em diferentes graus de risco em relação à diabetes. Quando esses algoritmos de aprendizagem de máquina têm acesso a classificação de instâncias anteriores, como uma base de dados de aprendizagem, eles são chamados de algoritmos de aprendizagem supervisionada (RUSSEL; NORVIG, 2013). Há diversos tipos de algoritmos de aprendizagem supervisionada:

- a) ***Nearest Neighbors (Vizinhos próximos)***: Em um espaço multidimensional, a classe de uma determinada instância provavelmente é similar a classe dos seus vizinhos. Entre os algoritmos dessa classe podemos encontrar o NearestCentroid, onde cada classe de instâncias é representada pelo centróide de seus membros, neste caso uma nova instância é predita com a classe com o centróide mais próximo; e o KNN, no qual a classificação de cada instância é calculada a partir de uma votação majoritária simples dos seus vizinhos mais próximos, e onde o número de vizinhos (K) é determinado pelo usuário.

Embora seja fácil de entender e implementar, o KNN tem como ponto desfavorável a necessidade de estabelecer empiricamente o melhor valor para K (SCIKIT-LEARN, 2022).

- b) **Máquinas de vetores de suporte (SVMs):** Esses algoritmos traçam uma linha (uma fronteira, um hiperplano) num espaço multidimensional que divide as diferentes classes de um problema. O hiperplano ótimo é escolhido entre o conjunto de N hiperplanos possíveis, como aquele cuja equidistância para a instância mais próxima de cada classe (vetores de suporte) é maximizada. Por ter diversas funções de Kernel, uma SVM é capaz de separar classes em problemas linearmente separáveis e problemas não-linearmente separáveis (OLIVEIRA JUNIOR, 2010).
- c) **Classificadores *Naive Bayes*:** Algoritmos classificadores que usam como base a probabilística. Eles são baseados no teorema de Bayes, calculando individualmente a probabilidade de que uma instância seja de cada classe, e na suposição "*Naive*" de independência condicional entre cada par de características dado o valor da variável classe (apesar dessa suposição de independência condicional não ser inteiramente verdadeira) (OLIVEIRA; MENDONÇA, 2000). No fim, a classe atribuída à instância é aquela cujo valor de probabilidade se mostrou maior. No geral, são algoritmos que requerem uma pequena quantidade de dados de treinamento para estimar os parâmetros necessários. Tem como Prós o fato de gastar pouco recurso computacional, não sofrer grandes influências de *outliers*, além de funcionar em problemas não lineares. Tem como contras o fato de supor que todos os atributos da instância têm a mesma relevância estatística. Uma outra característica que limita o uso algoritmos *Naive bayes* é que fato dele possuir como um de seus pressupostos a necessidade de independência entre os atributos das instâncias.
- d) **Árvores de decisão:** Esses algoritmos constroem conjuntos de nós que se ligam através de uma hierarquia, criando assim um grafo em forma de árvore. São algoritmos cujo entendimento é relativamente simples, pois sua representação é facilmente mostrada de forma similar a um fluxograma. Em cada nó da árvore são tomadas decisões, até que se atinjam os nós das folhas, que não tomam decisões, mas apenas armazenam o rótulo de uma determinada classe após uma série de decisões terem sido tomadas em nós anteriores. Cada uma dessas decisões é definida com base na entropia gerada ou no ganho de informação. Além de serem facilmente interpretáveis, as árvores lidam bem com problemas lineares e não lineares, além de lidar bem com dados não normalizados. Apesar disso, não geram bons resultados em pequenos conjuntos de dados, e às vezes é

necessária a construção de diversas árvores para obtenção de um bom resultado. Isso é feito, por exemplo, com o uso de algoritmos de florestas aleatórias (*Random Forests*), que utilizam uma fração das instâncias do conjunto total, para gerar múltiplas árvores cada uma delas gerando uma predição. Assim, uma alteração que poderia impactar uma determinada árvore, e assim alterar a predição do algoritmo, terá menos impacto, pois haverá centenas ou milhares de árvores de decisão trabalhando em conjunto (GÉRON, 2019). Essa ideia, de pegar os múltiplos de múltiplos classificadores mais simples para obter um resultado é o que está implementado nos métodos “*ensemble*”.

- e) **Redes Neurais Artificiais:** São algoritmos que tentam simular o comportamento de um conjunto de neurônios biológicos dos animais. Criados na década de 1940, esses algoritmos tiveram períodos de destaque, com grande quantidade de estudiosos focados no seu desenvolvimento, e também períodos em que foram praticamente abandonados. Uma rede neural artificial possui uma camada de entrada e uma de saída, entre elas há um número N de camadas escondidas, repletas de neurônios artificiais que se comunicam entre si. Nos dias atuais esses algoritmos vêm tendo muito destaque, principalmente devido ao surgimento de algoritmos com diversas camadas de neurônios ocultos, e que são capazes de realizar aprendizagem profunda, as chamadas redes neurais profundas (RNP), que muitas vezes superam outras técnicas de aprendizagem de máquina em problemas grandes e complexos. (GÉRON, 2019).

Atualmente centenas de artigos vêm sendo publicados envolvendo algum aspecto de aprendizagem de máquina e inteligência artificial para algum aspecto do diagnóstico, tratamento e assistência inteligente para a diabetes. Além do KNN, SVM, Naive Bayes, das árvores de decisão e redes neurais, os estudos também vêm utilizando métodos que unem diversos tipos de algoritmos para chegar a um melhor resultado, os chamados métodos “*ensemble*”, e métodos que seguem a lógica Fuzzy. Outros sistemas que utilizam técnicas como processamento de linguagem natural, além das técnicas clássicas de aprendizagem de máquina, também vêm sendo desenvolvidos com foco em assistentes inteligentes de interação mais simples com o ser humano portador da diabetes (CHAKI *et al.*, 2022)

2.7 DESENVOLVIMENTO DE BIOMARCADORES DIGITAIS

A ciência já avançou a um ponto em que diversas doenças possuem biomarcadores. Esses biomarcadores podem ser alguma medida fisiológica, que, quando alterada por indicar uma doença, ou mesmo mutações em genes que causam alteração na conformação ou expressão

de uma proteína. Hoje, por exemplo, estão se tornando comuns diversos biomarcadores genéticos para o diagnóstico de um dos grupos de doença mais estudadas, o câncer, embora muitos deles ainda não sejam tão amplamente utilizados devido a sua complexidade (KONCINA *et al.*, 2020). As tecnologias da informação vêm cada vez mais somando esforços na descoberta desses biomarcadores. Usos de métodos avançados de bioinformática, por exemplo, estão possibilitando a varredura completa do transcriptoma de um grupo de indivíduos, para descoberta de potenciais padrões que podem indicar a presença do câncer de mama (GYŐRFFY, 2021). A descoberta de biomarcadores, no entanto, vem se tornando um campo bem mais amplo. O uso de redes neurais profundas, por exemplo, está auxiliando a descoberta de biomarcadores a partir de imagens de doenças, expandindo o número de possíveis biomarcadores clínicos a partir de informações digitais (ECHLE *et al.*, 2021).

A popularização de dispositivos vestíveis² com foco em saúde, vem expandindo de forma extraordinária a possibilidade de descoberta de biomarcadores, já que milhares de dados são extraídos de um único indivíduo, geralmente de forma contínua. Esses biomarcadores são, no entanto, geralmente bem difíceis de serem desvendados, geralmente passando por alguma etapa de aplicação de modelos de aprendizagem de máquina e sendo conhecidos como biomarcadores digitais. Alguns tipos de relógios digitais possuem alta qualidade na captação de dados fisiológicos e por isso já estão autorizados, por algumas agências regulatórias, para o acompanhamento de doenças, como epilepsia. No caso específico desta doença, foram descobertos padrões de parâmetros fisiológicos (dados de atividade eletrodérmica e dados de movimentação) que estão ligados a episódios de convulsão e funcionam como marcadores para emissão de alertas ao indivíduo ou a um acompanhante (ONORATI *et al.*, 2017; PICARD; BOYER, 2021).

Geralmente dados utilizados como biomarcadores digitais incluem diversos dados derivados da fotopletismografia³, como a oximetria de pulso e frequência cardíaca; dados

² São dispositivos que estão diretamente acoplados no exterior do corpo humano, ou em suas roupas. Geralmente são materiais flexíveis e adaptáveis ao corpo humano, e que podem operar com baixo custo de energia, o que facilita sua captação de dados contínua. Roupas com tecidos especiais, tatuagens inteligentes, lentes de contato inteligentes, *smart rings* e *smartwatches* são exemplos de dispositivos vestíveis (IQBAL *et al.*, 2021).

³ A fotopletismografia é uma técnica que utiliza a óptica para detectar as alterações no fluxo sanguíneo dos tecidos, sendo capaz de acompanhar a pulsação derivada do batimento cardíaco e utilizada em dispositivos como oxímetros. Ela também pode ser utilizada para realizar o acompanhamento da pressão arterial e da atividade cardíaca para avaliar alguns tipos de doenças vasculares periféricas (BISWAS *et al.*, 2019).

derivados de actígrafos⁴; dados de atividade eletrodérmica, assim como temperatura da pele, exposição à radiação (DUNN; RUNGE; SNYDER, 2018). A frequência cardíaca diz respeito ao número de batimentos cardíacos por minuto. Em geral, uma pessoa em repouso tem entre 60 e 100 batimentos cardíacos. Esse ritmo é aferido por medidores especializados, que tem muita acurácia, mas também pode ser medido por relógios ou pulseiras inteligentes através da fotopletismografia. Embora o grau de acurácia da medição seja um pouco menor nesses últimos equipamentos, eles são amplamente dispersos entre a população, ao contrário dos medidores especializados em forma de faixa que envolve todo o peito do indivíduo. Nestes aparelhos vestíveis em forma de relógios ou pulseiras inteligentes, o batimento cardíaco é medido através do fluxo sanguíneo presente abaixo da pele (CORLISS, 2021). Quando há relaxamento dos vasos sanguíneos há mais absorbância da luz verde (emitida pelo aparelho), já quando há contração a absorbância é menor. Esse movimento dos vasos afeta o volume sanguíneo, que é o responsável pela diferença de leitura. A informação desse volume sanguíneo no pulso (*blood volume pulse*) também se torna uma métrica estimada pelos relógios inteligentes (PICARD; BOYER, 2021).

Outra métrica envolvendo os batimentos cardíacos, que também é extraída pelos relógios inteligentes é o intervalo interbatidas, também conhecido por intervalo RR, que é o intervalo entre dois picos R sucessivos. Os picos R são um dos 5 picos derivados da atividade elétrica do coração. Um eletrocardiograma genérico de um paciente saudável possui uma onda P, um complexo QRS e uma onda T. O complexo QRS é derivado dos potenciais elétricos gerados durante a despolarização dos ventrículos antes do episódio de contração. Alterações no intervalo entre esses picos podem representar variações normais dos batimentos derivados de diferentes respostas fisiológicas, mas quando intensificadas, podem indicar problemas de saúde mas graves, como arritmia cardíaca (GUYTON; HALL, 2006; PAULO; BOCK, 2014).

Uma outra métrica extraída a partir de relógios e pulseiras inteligentes é a atividade eletrodérmica (EDA). Durante muito tempo a EDA foi utilizada apenas para aferição de resposta ao estresse num indivíduo, a partir do cálculo da média de seus valores no intervalo de tempo avaliado. Essa métrica avalia a condutância da superfície da epiderme, ou seja, fenômenos elétricos que estão ocorrendo na superfície da pele, e que são derivados da alteração

⁴ São aparelhos que registram continuamente a movimentação do indivíduo e são geralmente utilizados para avaliação do sono. Eles possuem sensores de acelerometria triaxial, ou seja, acelerômetros que medem movimentos nas três dimensões (TELLES *et al.*, 2001).

da atividade das glândulas sudoríparas. Isto é uma forma indireta de medir a atividade do sistema nervoso simpático. Hoje a EDA também vem sendo bastante utilizada para avaliar características patológicas de um indivíduo, além de refletir fenômenos como picos de estresse, perturbações de sono, alterações na sudorese (PICARD; BOYER, 2021; POSADA-QUINTERO; CHON, 2020).

2.8 TRABALHOS RECENTES RELACIONADOS

O desenvolvimento de dispositivos vestíveis, que podem extrair uma série de métricas de saúde do indivíduo, fez surgir uma imensa fonte de dados fisiológicos de pacientes, e como visto anteriormente, através da aplicação de algoritmos de aprendizagem de máquina esses dados podem ser associados a variação glicêmica. Nos últimos 5 anos, a busca por um dispositivo vestível que possa, eficazmente, substituir aparelhos tecnológicos não invasivos no diagnóstico da curva glicêmica de um indivíduo, tem se intensificado

Além da utilização de dispositivos vestíveis de uso comum, os estudos têm desenvolvido dispositivos que estão conectados com a *Internet*, armazenando dados em nuvem a todo tempo, podendo fornecer informações a um paciente e seu médico em tempo real. Esse contexto em que dispositivos médicos estão cada vez mais conectados com servidores em nuvem, e sem a necessidade de fios, é chamado de Internet das Coisas Médicas (*Internet of Medical Things - IoMT*). Um campo com intenso desenvolvimento de *hardware*, *software* e busca por biomarcadores de doenças (QURESHI; KRISHNAN, 2018).

Um dos dispositivos inseridos no contexto do IoMT, o iGLU, foi criado com a intenção de verificar a glicose sérica de indivíduos portadores de diabetes, pré-diabetes ou de indivíduos saudáveis. Como inovações ele trouxe uma predição de glicose baseada em espectroscopia de absorbância e refletância com o uso de NIR. As leituras dos sensores eram processadas por uma rede neural profunda (RNP) com 10 camadas ocultas, gerando uma alta correlação entre a glicose real e a predita ($R=0,953$) (JAIN; JOSHI; MOHANTY, 2020). Uma atualização do sistema, o iGLU 2.0, foi criada logo em seguida. Utilizando uma tecnologia de *hardware* similar e um melhor modelo de ML, baseado em uma regressão polinomial múltipla de grau 3, a nova versão foi capaz de estimar glicemias na faixa de 80-420 mg/dL, mostrando um coeficiente de correlação ($R=0,97\%$), entre a glicose real e a estimada, superior ao iGLU original (JOSHI *et al.*, 2020). Apesar do alto poder preditivo do iGLU, ele ainda está na fase de protótipo e seu design não provê uma boa experiência para o usuário, uma vez que necessita que ele fique com

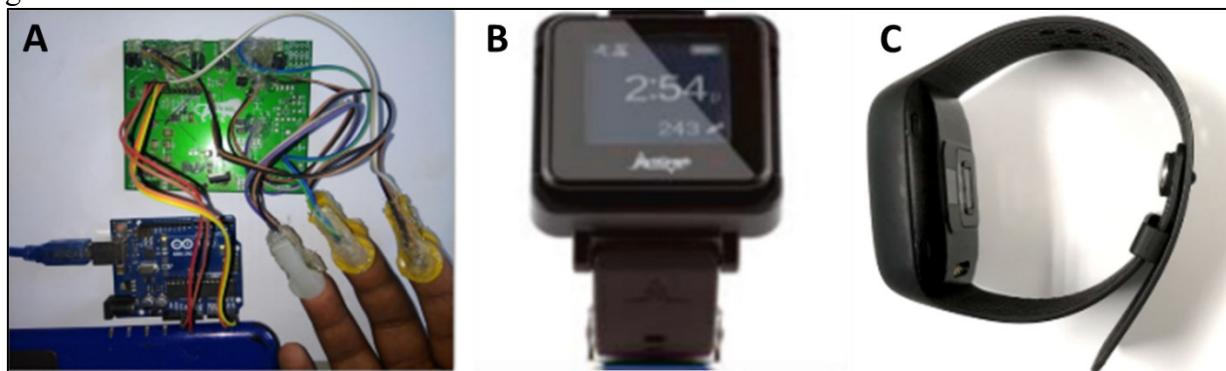
3 dedos ligados a sensores (Fig. 5A), isso dificulta sua popularização como dispositivo de uso cotidiano. Além do poder preditivo alto, as pesquisas também vêm focando em sensores que também utilizam o método de NIR porém com um menor consumo de energia, o que proporciona um maior tempo de monitoramento antes do fim da bateria do dispositivo (IBRAHIM; DARWEESH; SOLTAN, 2022).

Uma alternativa ao desenvolvimento de sensores específicos de glicose seria a utilização de aparelhos já existentes, como é o caso do actígrafo (sensor de actimetria), que é muito similar a um relógio de pulso. A adequação dos dados captados por este sensor, para predição de diabetes foi avaliada por Ramazi *et al.* (2021). Os autores deste estudo também coletaram dados da glicemia de indivíduos através de um aparelho de monitoramento contínuo de glicose. Dados demográficos (idade e sexo), dados físicos (peso, altura, circunferência da cintura), dados de exames sanguíneos (taxas de colesterol, triglicerídeos e hemoglobina glicada) e de exames de urina também foram obtidos. A partir disso, os dados de glicose e actimetria foram inicialmente convertidos por uma transformada de Fourier, o que ajudava no reconhecimento de padrões. Estes então foram passados para uma rede neural convolucional (CNN) responsável por analisar as dependências e interações entre os sensores. A saída desta CNN atuou como input de uma rede neural GRU (*gated recurrent unit*). O resultado da GRU foi processado paralelamente com os dados físicos, demográficos e sanguíneos de forma que todo esse *pipeline* construído pelos autores pudesse realizar tarefas de regressão e classificação. Como tarefa de classificação, o pipeline construído prevê, com acurácia de 0,9, 0,92, 0,87 e 0,88 se os dados de HbA1c, triglicerídeos, colesterol HDL e LDL, respectivamente, irão aumentar ou diminuir um ano após a coleta dos dados de um indivíduo. Como tarefa de regressão o pipeline prevê, com erros relativamente pequenos ($RMSE < 18$), o valor dessa diferença das taxas, entre o dado atual e o predito para um ano depois.

Além da actimetria, os dados de outros tipos de sensores também vêm sendo relacionados com a diabetes, como os sensores de fotopletismografia. Prabha *et al.* (2021a) utilizaram dados de fotopletismografia de dispositivos vestíveis em 217 indivíduos para tentar predizer a presença da doença. Eles pré-processaram os dados contínuos da fotopletismografia, em janelas de 5 segundos e utilizaram técnicas de processamento de som (*Mel-frequency cepstral coefficients* - MFCC), seguido de aplicação de KNN e SVM para classificar o indivíduo em saudável, pré-diabético ou diabético. Seus resultados mostraram que fotopletismografia pode ser uma boa preditora para diabetes, uma vez que a acurácia do KNN foi de 83,87 % e do SVM+PCA foi de 92,28%. Os mesmos autores mostraram posteriormente que seu método

poderia chegar a acurácia de 99,74%, com o uso de um algoritmo de ML chamado de *extreme gradient boost* (XGBoost) (PRABHA *et al.*, 2021b). Além da presença da doença, a fotopletismografia também pode ser utilizada para estimativa da glicemia de uma pessoa. Hina e Saadeh (2020) mostraram que o uso de SVM com um kernel gaussiano acoplado a um processador digital para eliminação de ruídos é capaz de gerar uma glicemia predita com uma alta correlação ($R=0,968$) com a glicemia real, assim como um baixo RMSE (11,20). Apesar disso da alta correlação, o trabalho de Hina e Saadeh (2020) não utilizou dispositivos comercialmente disponíveis, o que limita a disseminação do método.

Figura 5 – Sensores não invasivos utilizados como fonte de dados em estudos de predição de glicemia.



Legenda: (A) Protótipo do iGLU, (B) sensor de actimetria e (C) *smartwatch* Empatica E4.

Fonte: Elaborado pelo autor com base em imagens publicadas por Jain, Joshi e Mohanty (2020), Ramazi *et al.* (2021) e Yin *et al.* (2021).

Outros dispositivos vestíveis vendidos comercialmente contam com um maior número de sensores que o dispositivo utilizado por Ramazi *et al.* (2021). Eles geralmente possuem sensores de eletrocardiograma, sensores de resposta galvânica da pele, sensores de fotopletismografia e acelerômetros. Em 2021 foi proposto o DiabDeep, um *framework* que combinava sinais diversos como: I) dados fisiológicos como ritmo cardíaco, temperatura corpórea, resposta galvânica da pele, e pulsação, que são justamente captados por estes dispositivos vestíveis que funcionam como *smartwatches* (Fig. 5C); II) dados demográficos, como idade, peso, gênero, altura, sexo, características que podem ser obtidas a partir da interação de um usuário com um aplicativo em *smartphone*. Esse sistema utilizava redes neurais treinadas a partir de um paradigma “*grow and prune*”, onde primeiro se permite que os neurônios criem diversas conexões e depois se elimina aquelas que não são significativas. O framework foi construído com dois tipos de rede, uma delas executada em um servidor remoto, e outra que é executada no dispositivo periférico que o indivíduo carrega, como seu *smartphone*. Enquanto a arquitetura da primeira focava em alta acurácia, a última é menos robusta, focando em eficiência uma vez

que precisa rodar em um ambiente com restrição de armazenamento e memória. Apesar disso, quando utilizadas em conjunto, essas redes foram capazes de prever com alta acurácia (>95%), se indivíduos eram ou não portadores de diabetes. Além disso, elas também foram capazes de classificar, com acurácia por volta de 95%, qual o tipo de diabetes que o indivíduo possuía, se a DM1 ou a DM2 (YIN *et al.*, 2021).

Bent *et al.*, (2021a) publicaram um estudo que também utilizava dados de sensores derivados da pulseira Empatica E4 em conjunto com dados demográficos dos pacientes. Além disso, eles adicionaram dados relacionados a hábitos comportamentais do indivíduo, como dados de atividade física, ritmo circadiano, estresse e dados de dieta envolvendo a ingestão de calorias, proteínas, açúcares e outros carboidratos. Para cada dado do sensor, eles extraíram 7 métricas estatísticas, resultando num conjunto de dados final com 69 características. Os autores utilizaram essas características para tentar prever a glicose intersticial e tentar classificar os níveis de glicose em alta, baixa ou normal, com limiares que variavam para cada indivíduo. Para a tarefa de classificação os autores utilizaram redução de dimensionalidade e um algoritmo de árvore de decisão, obtendo uma acurácia de 84,3%. Para a tarefa de predição eles utilizaram o algoritmo XGBoost obtendo um RMSE de 21,22 mg/dL, o que é um erro relativamente alto. Os autores também observaram que os dados mais importantes para o seu modelo eram dados relacionados com a dieta do indivíduo. Dados desse tipo, no entanto, não são coletados automaticamente por pulseiras ou outro tipo de aparelho, exigindo que o usuário preencha sobre seus horários e tipo de alimento nas refeições.

Além da tarefa de classificação do tipo de diabetes a partir de sinais fisiológicos capturados por *smartwatches*, estudos também buscam realizar uma tarefa de regressão para conseguir prever índices como a HbA1C e algumas métricas de glicemia. Um estudo com 8 pessoas pré-diabéticas e outras 8 com glicemia em níveis normais mostrou que algumas métricas, como a média da glicose diária de um paciente, podem ser estimadas, com baixo nível de erro, a partir de um conjunto de sinais fisiológicos capturados pelos *smartwatches*. Neste estudo foram construídas florestas aleatórias específicas para cada uma das 27 métricas de variação de glicose avaliadas, das quais 11 foram obtidas com alta acurácia. Da mesma forma, a HbA1c também pode ser estimada, com baixo erro (RMSE = 0,281), a partir do mesmo conjunto de sinais fisiológicos captados por sensores de acelerometria, frequência cardíaca, atividade eletrodérmica ou temperatura. Apesar disso, este estudo não contou muita diversidade étnica e racial, o que pode limitar seus resultados para a população em geral (BENT *et al.*, 2021b).

Um outro estudo também focou num público composto por paciente diabéticos e não-diabéticos, mas utilizando dados fisiológicos obtidos de um tipo diferente de *smartwatch* que o utilizado por Bent *et al.* (2021b). Além da frequência cardíaca e atividade eletrodérmica, outros dados como, nível de oxigênio no sangue (SPO2), pressão arterial diastólica, pressão arterial sistólica e temperatura corporal também foram utilizados por (AHMED *et al.*, 2023) para estimar os níveis de glicose em pacientes diabéticos. Neste estudo, Ahmed *et al.* (2023) testaram os algoritmos SVM e Random Forest, uma RNP, e um algoritmo com a lógica Fuzzy, para predição da glicemia. Os 4 algoritmos testados resultaram em baixos valores de RMSE, que variou entre 0,183 até 0,197, valores menores que o obtido por Bent *et al.* (2021b). Interessantemente, o algoritmo de Random Forest superou os algoritmos de redes neurais profundas na predição da glicemia em relação a métrica de RMSE (RMSE = 0,183), que descreve o desvio médio dos valores reais de glicemia em relação aos valores preditos pelo algoritmo.

Dados como frequência cardíaca, atividade eletrodérmica e movimentação também foram recentemente utilizados como informações de entrada para uma árvore de decisão que mostrou uma certa capacidade (AUC = 0,76) de detectar eventos de hipoglicemia em pacientes diabéticos que faziam terapia com insulina e por isso tinham mais chances de desenvolver hipoglicemia. Para isso Lehmann *et al.*, (2023) utilizaram um relógio Garmin vivoactive 4S em conjunto com uma pulseira Empatica E4 e um glicosímetro Dexcom G6 para coletar dados de paciente durante 30 dias.

Apesar do fato de que trabalhos recentes vêm conseguindo correlacionar, com baixo erro, métricas de glicose com dados extraídos de sensores de *smartwatches*, uma das limitações para sua aplicação em larga escala é o fato de a maior parte das pulseiras e relógios inteligentes não possuírem todos os sensores que geralmente são utilizados nos estudos (BENT *et al.*, 2021b). Outro ponto fraco a ser levantado é o fato de alguns estudos apontarem que os dados gerados pelos sensores de relógios e pulseiras inteligentes são muito predispostos a conterem erros de leituras, artefatos e possuem muita interferência de ruído, e que por isso não devem ser fontes ótimas de dados para alimentar diversos tipos de algoritmos de aprendizagem de máquina. E que por isso devem ser filtrados e decompostos antes do uso em qualquer tarefa de predição ou classificação (FÖLL *et al.*, 2021).

Um resumo de todos os trabalhos relacionados revisados pode ser visto no Quadro 1.

Quadro 1 – Principais características de trabalhos relacionados.

Autor/data	Equipamento/ método de coleta de dados	Parâmetros avaliados	Algoritmos utilizados	Finalidade	Desempenho
(JAIN; JOSHI; MOHANTY, 2020)	iGLU/ espectroscopia de absorbância e refletância com o uso de NIR	Glicose Sérica	Rede Neural Profunda (10 neurônios e 10 camadas ocultas)	Predição do índice glicêmico	RMSE = 11,56 MAD = 9,89
(JOSHI <i>et al.</i> , 2020)	iGLU2/ espectroscopia de absorbância e refletância com o uso de NIR	Glicose Sérica	Régressão Polinomial Múltipla de grau 3	Predição do índice glicêmico	RMSE = 13,57 MAD = 9,42
Ramazi <i>et al.</i> (2021)	Actígrafo, exame sanguíneo, exame de urina, avaliação física, dados demográficos.	Movimentação em tempo real, idade, sexo, peso, altura, circunferência da cintura, colesterol, triglicerídeos, hemoglobina glicada	Rede Neural Convolucional (CNN) seguida de uma rede neural GRU	Classificação e predição de taxas como HbA1c, triglicerídeos, colesterol HDL e LDL	Classificação de HbA1C: • Acurácia = $0,90 \pm 0,02$ • AUROC = $0,88 \pm 0,02$ Predição de HbA1C: • RMSE = $1,37 \pm 0,8$
Prabha <i>et al.</i> (2021a)	Pulseira Empatica E4	Fotopletismografia	KNN e SVM	Classificação de pacientes saudáveis,	KNN: Acurácia = 83,87%

				diabéticos e pré-diabéticos.	SVM+PCA: Acurácia = 92,28%.
Hina e Saadeh (2020)	Protótipo que utiliza uma emissão de LED NIR com um comprimento de onda de $\lambda=940$ nm	Fotopletismografia	SVM	Predição do índice glicêmico	RMSE = 11,20 mARD = 7,62%
(YIN <i>et al.</i> , 2021)	Pulseira Empatica E4; questionários em smartphones Samsung Galaxy S4.	Ritmo cardíaco, temperatura corpórea, resposta galvânica da pele, pulsação, idade, peso, gênero, altura, sexo.	Redes Neurais	Classificar indivíduos portadores e não portadores de diabetes; diferenciar DM1 e DM2.	Acurácia = 94,6% para predição de três classes; Acurácia = 96,3% para predição de duas classes.
(BENT <i>et al.</i> , 2021a)	Pulseira Empatica E4 e Glicosímetro Dexcom G6	Acelerometria, frequência cardíaca, atividade eletrodérmica e temperatura, dados demográficos, dados de hábitos: estresse, dietas, exercícios, ritmo circadiano e outros.	Árvore de decisão e XGBoost	Classificar o nível de glicose e prever a glicose intersticial.	Acurácia = 82% Recall = 82% Precisão = 82,3% F1 score = 82,1%

(BENT <i>et al.</i> , 2021b)	Pulseira Empatica E4	Acelerometria, frequência cardíaca, atividade eletrodérmica e temperatura.	Random Forest	Estimar 27 métricas de variação de glicose e a HbA1C em pacientes saudáveis e pré-diabéticos.	<p>HbA1C:</p> <ul style="list-style-type: none"> • RMSE= 0,357 • MAPE= 5,1% <p>Métricas restantes:</p> <ul style="list-style-type: none"> • $0,25 < \text{RMSE} < 763$ • $4,14 < \text{MAPE} < 162,79$
Ahmed <i>et al.</i> (2023)	Pulseira RiverSong Wave O2	Frequência cardíaca, atividade eletrodérmica, nível de oxigênio no sangue (SPO2), pressão arterial diastólica, pressão arterial sistólica, temperatura corporal e tremores.	SVM, Random Forest, RNP e lógica Fuzzy.	Predição do índice glicêmico	<p>Random Forest:</p> <ul style="list-style-type: none"> • $0,183 < \text{RMSE} < 0,189$ • MAE = 0,102
(LEHMANN <i>et al.</i> , 2023)	Relógio Garmin vivoactive 4S e Pulseira Empatica E4	Frequência cardíaca, atividade eletrodérmica e movimentação.	Árvore de decisão.	Predição de episódios de hipoglicemia (tarefa de classificação).	AUROC = $0,76 \pm 0,07$

3 OBJETIVOS

3.1 OBJETIVO GERAL

Avaliar modelos de aprendizagem de máquina a partir de dados de dispositivos vestíveis e verificar se eles são capazes de prever, com eficácia, se um indivíduo pode ser caracterizado como pré-diabético.

3.2 OBJETIVOS ESPECÍFICOS

- a) Obter e realizar o processamento dos dados brutos da pulseira Empatica E4 obtidos do banco de dados PhysioNet;
- b) Realizar um teste de prova de conceito para avaliar o desempenho de vários algoritmos na tarefa de classificação do estado de saúde de indivíduos com base nas leituras derivadas da pulseira Empatica E4;
- c) Realizar um teste de prova de conceito para avaliar a performance de diversos algoritmos na tarefa de classificação do estado de saúde de indivíduos a partir de uma base contendo atributos engenheirados derivados da pulseira Empatica E4.

4 METODOLOGIA

4.1 FONTE DE DADOS

Os dados foram obtidos a partir do banco de dados PhysioNet (GOLDBERGER *et al.*, 2000), um portal que armazena diversas bases de dados que envolvem experimentos sobre a fisiologia humana. O banco de dados utilizado foi o “*BIG IDEAs Lab Glycemic Variability and Wearable Device Data*” (<https://physionet.org/content/big-ideas-glycemic-wearable/1.0.0/>) publicado inicialmente em setembro de 2022 e atualizado em março de 2023 (CHO *et al.*, 2023).

4.2 DESCRIÇÃO DO BANCO DE DADOS

O banco de dados possui informações sobre dados fisiológicos de 16 indivíduos. Desses, 9 eram do sexo feminino e 7 eram do sexo masculino. Metade dos indivíduos possuíam valores de HbA1C superiores a 5,7 e dessa forma são portadores de pré-diabetes (5 mulheres e 3 homens). A outra metade possuía níveis normais de HbA1C (4 homens e 4 mulheres). Todos os indivíduos possuíam idades que variavam entre 35 e 65 anos.

Os dados disponibilizados para cada indivíduo são dados de leitura de métricas fisiológicas extraídas a partir da pulseira Empatica E4 e a partir do glicosímetro Dexcom 6. A pulseira Empática forneceu os dados de atividade eletrodérmica, volume sanguíneo no pulso (*blood volume pulse*), ritmo cardíaco, intervalo interbatidas (intervalo entre os batimentos cardíacos), temperatura da pele e dados de movimentação (acelerômetro nos eixos X, Y e Z). Cada um desses dados foi coletado simultaneamente, porém o intervalo de tempo entre as leituras variava de acordo com o tipo de sensor. Os dados de movimentação eram lidos 32 vezes por segundo, enquanto os dados de pulsação eram lidos no dobro dessas vezes (64 leituras por segundo). Outros dados têm leituras menos frequentes, como a atividade eletrodérmica e a temperatura da pele, ambas aferidas 4 vezes por segundo; o intervalo interbatidas a cada 781 milésimos de segundo, e por último o ritmo cardíaco, uma vez por segundo. O glicosímetro Dexcom 6 forneceu dados diversos, mas principalmente dados da glicemia a cada intervalo de 3 minutos. Os dados do Dexcom, no entanto, não foram utilizados neste estudo.

4.3 PRÉ-PROCESSAMENTO DOS DADOS

4.3.1 Pré-processamento para construção de uma base de dados de sinais brutos

O primeiro passo do pré-processamento foi padronizar o intervalo de tempo de leitura para cada sensor. Isso foi realizado considerando-se a média das leituras numa janela de tempo de 1 minuto, ou seja, todas as leituras obtidas no intervalo de um minuto foram somadas e divididas pelo número de leituras. Assim, cada sensor passou a apresentar apenas um valor por minuto. Essa transformação foi realizada com o método “*resample*” da biblioteca Pandas (versão 1.4.3) do Python (versão 3.9.12).

Assim como os dados originais da base, os dados pré-processados da base estavam separados por indivíduo e por sensor utilizado para sua captação. Dessa forma havia 16 arquivos para cada sensor, cada um possuindo os dados de um determinado indivíduo. O passo seguinte foi unir todos os dados de um indivíduo, neste momento foram considerados os seguintes dados: I) atividade eletrodérmica (EDA), II) ritmo cardíaco, III) intervalo interbatidas (IBI - intervalo entre os batimentos cardíacos), IV) temperatura da pele, V) pulsação (*blood volume pulse*), VI) sexo e VII) categoria de saúde. Excluindo-se os dados VI e VII, todos os outros dados variaram com o passar dos minutos e por isso o conteúdo dos diversos arquivos foram mesclados através de um procedimento de “*full outer join*” utilizando como chave primária o momento de cada leitura. A descrição dos atributos da base de dados final está mostrada no Quadro 2.

Quadro 2 – Descrição dos atributos da base de dados

Atributo	Descrição	Tipo
Coluna 0	Atividade eletrodérmica	Variável Real
Coluna 1	Ritmo cardíaco	Variável Real
Coluna 2	Intervalo interbatidas	Variável Real
Coluna 3	Volume sanguíneo no pulso	Variável Real
Coluna 4	Temperatura da pele	Variável Real
Coluna 5	Sexo	Variável categórica (binária) 0 = Feminino 1 = Masculino
Coluna 6	Categoria de saúde de onde os dados foram coletados (Classe)	Variável categórica (binária) 0 = Hb1AC < 5,6 (Indivíduo Saudável) 1 = Hb1AC ≥ 5,7 (Indivíduo Pré-diabético)

Como os dados são reais ou binários, não houve nenhuma necessidade de transformação das categorias. No entanto, apesar de não necessitarem desse tipo de transformação, os dados da base não estavam normalizados. Para realização desse procedimento foi utilizada a função StandardScaler() do pacote Scikit-Learn (PEDREGOSA *et al.*, 2011), que transforma um valor X em um valor Y a partir da função: $Y = (X - U) / S$, onde U é a média dos valores de um atributo e S é o desvio padrão.

Dados faltantes presentes na base foram tratados a partir de imputação considerando o método dos vizinhos mais próximos. Nesse método, os valores faltantes são imputados utilizando os valores, do mesmo atributo, das instâncias com características mais similares as instâncias cujo valores estão ausentes. Para realização da imputação foi utilizada a classe KNNImputer do pacote Scikit-Learn (versão 1.1.1). No caso de uma determinada instância conter apenas valores faltantes, não foi realizada uma imputação, mas sim a exclusão da instância da base de dados.

4.3.2 Pré-processamento para construção de uma base de dados de sinais decompostos

Para avaliar se há diferenças do desempenho da tarefa de classificação entre a utilização de dados brutos derivados da pulseira Empatica E4 e a utilização de dados frutos da decomposição de sinais, produzimos uma segunda base de dados. Essa base era composta pelos dados de ritmo cardíaco, temperatura da pele, volume sanguíneo do pulso, sexo e pela decomposição de dois sinais, o intervalo interbatidas e a atividade eletrodérmica. O processo de engenharia de atributos desses dois sinais foi realizado através da biblioteca FLIRT (versão 0.0.2) (FÖLL *et al.*, 2021). Essa biblioteca possui diversas funções, dentre elas funções específicas para lidar com os dados produzidos pela pulseira Empatica E4.

Para o processamento dos dados foram utilizadas duas funções da biblioteca, a função `flirt.hrv.get_hrv_features()` e `flirt.eda.get_eda_features()`, aplicadas nos dados de IBI e EDA respectivamente. A primeira função recebeu como parâmetros: `window_length = 1200`, `window_step_size = 60`, `domains = ['td', 'fd', 'nl', 'stat']` e `threshold = 0,5`. O parâmetro `window_step_size` diz respeito ao nível de granularidade da análise em relação ao tempo, um valor igual a 60 (como configurado) quer dizer que o processamento será feito a cada intervalo de 60 segundos, de forma a estar na mesma escala de tempo que o restante dos dados dos outros três sensores. A opção de utilizar a lista `['td', 'fd', 'nl', 'stat']` no parâmetro `'domains'` gera características em todos os domínios possíveis para esse tipo de dado, como o domínio de tempo

e frequência, além de características estatísticas e características não-lineares capazes de quantificar a incerteza nas sequências dos intervalos interbatidas. Essa função foi responsável por decompor o dado de intervalo interbatidas em 54 novos atributos. Já a segunda função recebeu como parâmetros: *window_length*= 1200, *window_step_size* = 60, produzindo ao final 44 atributos, dos quais 2 (*phasic and tonic entropy*) foram eliminados por conter muitos valores infinitos para diversas instâncias.

As etapas subsequentes de normalização e imputação de dados foram realizadas da mesma forma que o procedimento descrito anteriormente. Porém, instâncias onde havia valores faltantes para os dados decompostos pelo FLIRT foram excluídas do conjunto de dados.

4.4 ALGORITMOS AVALIADOS

Objetiva-se saber se há uma diferença nos parâmetros fisiológicos, captados pela pulseira Empatica E4, que seja suficiente para classificar um indivíduo em saudável ou pré-diabético. Para esta tarefa de classificação utilizamos a biblioteca Scikit-Learn, que possui todas as funções necessárias para cada uma das etapas de um projeto envolvendo aprendizagem de máquina.

A fim de encontrar um algoritmo com maior capacidade preditora testamos três classes distintas de algoritmos: máquinas de vetores de suporte (SVM), implementada pelas classes *svm.SVC()* e *svm.LinearSVC()*; árvores de decisão, implementada pelas classes *ensemble.RandomForestClassifier()* e *tree.DecisionTreeClassifier()*; e redes neurais implementada pela classe *neural_network.MLPClassifier()*; assim como um meta-estimador das classe *ensemble.StackingClassifier()* construído a partir das melhores configurações dos melhores algoritmos de cada tipo de algoritmo.

Todos os algoritmos de aprendizagem supervisionada avaliados neste trabalho tiveram seus hiperparâmetros otimizados através de várias permutações entre os valores dos parâmetros, realizadas em conjunto com uma validação tipo *10-fold cross validation* (estratificado). Isso foi realizado a partir da função *GridSearchCV*, que recebe um objeto do tipo pipeline, um o valor *k* (neste caso *k*=10) para a validação cruzada e um conjunto de parâmetros a serem testados. Os parâmetros utilizados para a otimização de cada um dos algoritmos testados estão listados no Quadro 3.

Quadro 3 – Hiperparâmetros avaliados para cada um dos algoritmos de aprendizagem de máquina.

Algoritmo	Hiperparâmetros: valores
SVC	C: [1, 0.1]
	gamma: ['auto','scale']
	kernel: ['rbf', 'sigmoid']
LinearSVC	C: [1, 0.5, 0.1, 0.05]
	max_iter: 3000
	loss: ['squared_hinge','hinge']
DecisionTreeClassifier	criterion: ['gini', 'entropy']
	min_samples_leaf: [2,3]
	max_depth: [3,5,6,7,9,11,13,15,17,19]
	min_samples_split= [2,3,5,7]
Random ForestClassifier	criterion: ['gini', 'entropy']
	min_samples_leaf: [1,2,3,4]
	max_depth: [5, 10, 20, 30, 50, None]
	min_samples_split= [2,3,5]
MLPClassifier	hidden_layer_sizes: [(10,10), (10,20), (20, 50), (20,20,30), (50,100, 50), (20, 50, 50), (20,20,30), (10,10,10,10)]
	alpha: [0.00001, 0.000001, 0.0001]
	learning_rate_init: [0.001, 0.0005]
	activation: ['tanh', 'relu']

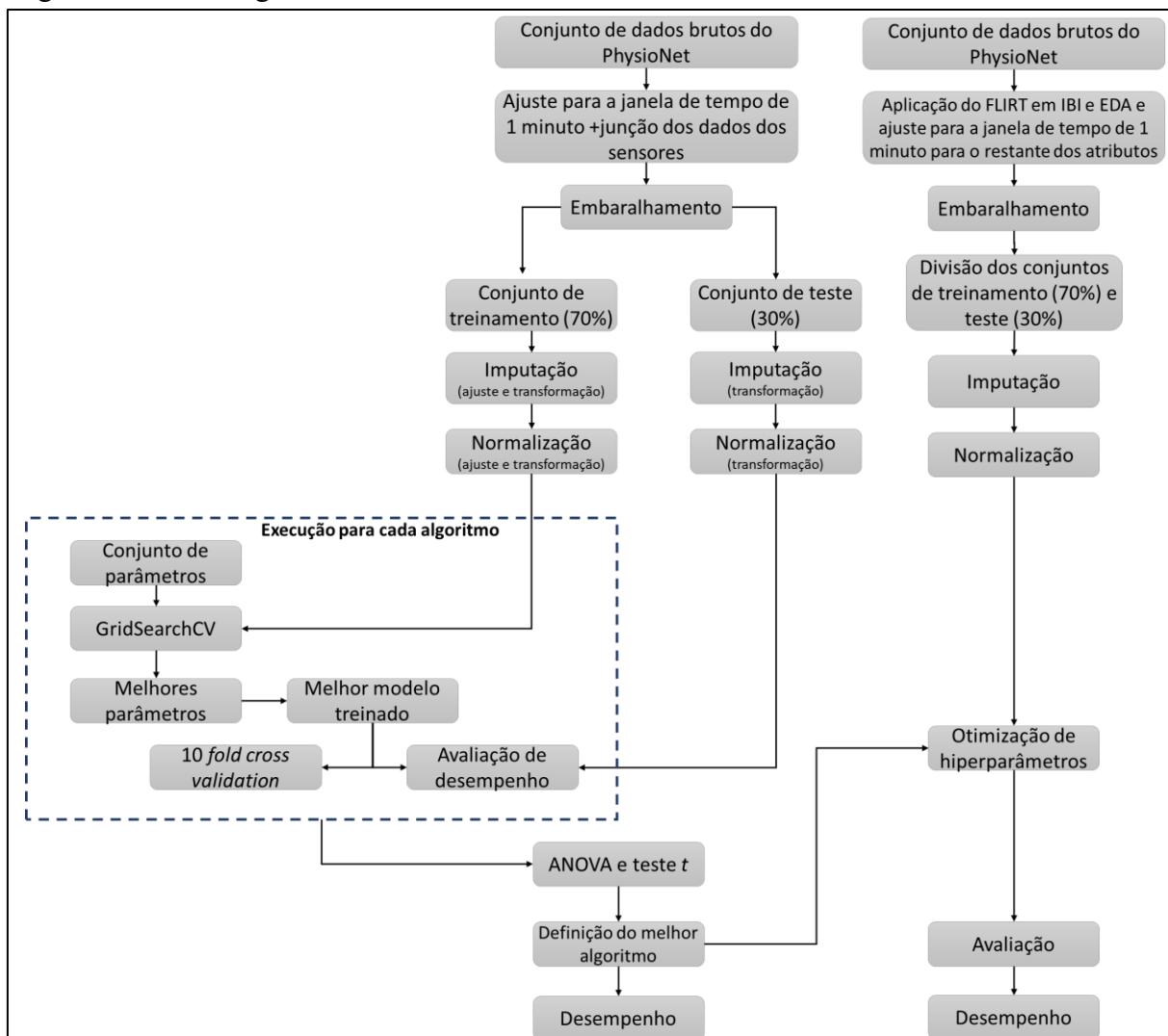
4.5 VALIDAÇÃO ESTATÍSTICA

Após a obtenção dos melhores hiperparâmetros para cada algoritmo, os dados foram divididos em conjuntos de treinamento (70% da base) e de teste (30% da base), sendo obtidas as métricas de Acurácia, Precisão, Sensibilidade e Área sob a Curva ROC. Essas mesmas métricas foram obtidas a partir da execução de um uma validação cruzada *10-fold* no conjunto de treinamento. Obtendo-se então 10 valores de cada métrica para cada algoritmo no conjunto de treinamento e 1 valor para a comparação de treinamento e teste. Foi realizado um teste de ANOVA para verificar se algum dos algoritmos testados é capaz de superar os outros em relação a cada uma das 4 métricas avaliadas. Um teste *t* de *student* foi realizado entre o algoritmo com o maior valor médio de uma determinada métrica e o valor médio dela para cada

um dos algoritmos restantes, de forma a verificar-se quais de fato eram os melhores algoritmos de predição. O tempo de execução de cada algoritmo para as funções de treino e teste também foi calculado, para ser utilizado como critério de desempate quando os scores de qualidade dos métodos não diferirem significativamente.

A visualização compacta das etapas realizadas neste trabalho está mostrada na Figura 6.

Figura 6 – Resumo geral do fluxo de trabalho utilizado neste estudo.



Fonte: Elaborado pelo Autor

5 RESULTADOS E DISCUSSÃO

5.1 ANÁLISE EXPLORATÓRIA DA BASE CONTENDO OS DADOS BRUTOS DERIVADOS DOS SENsoRES

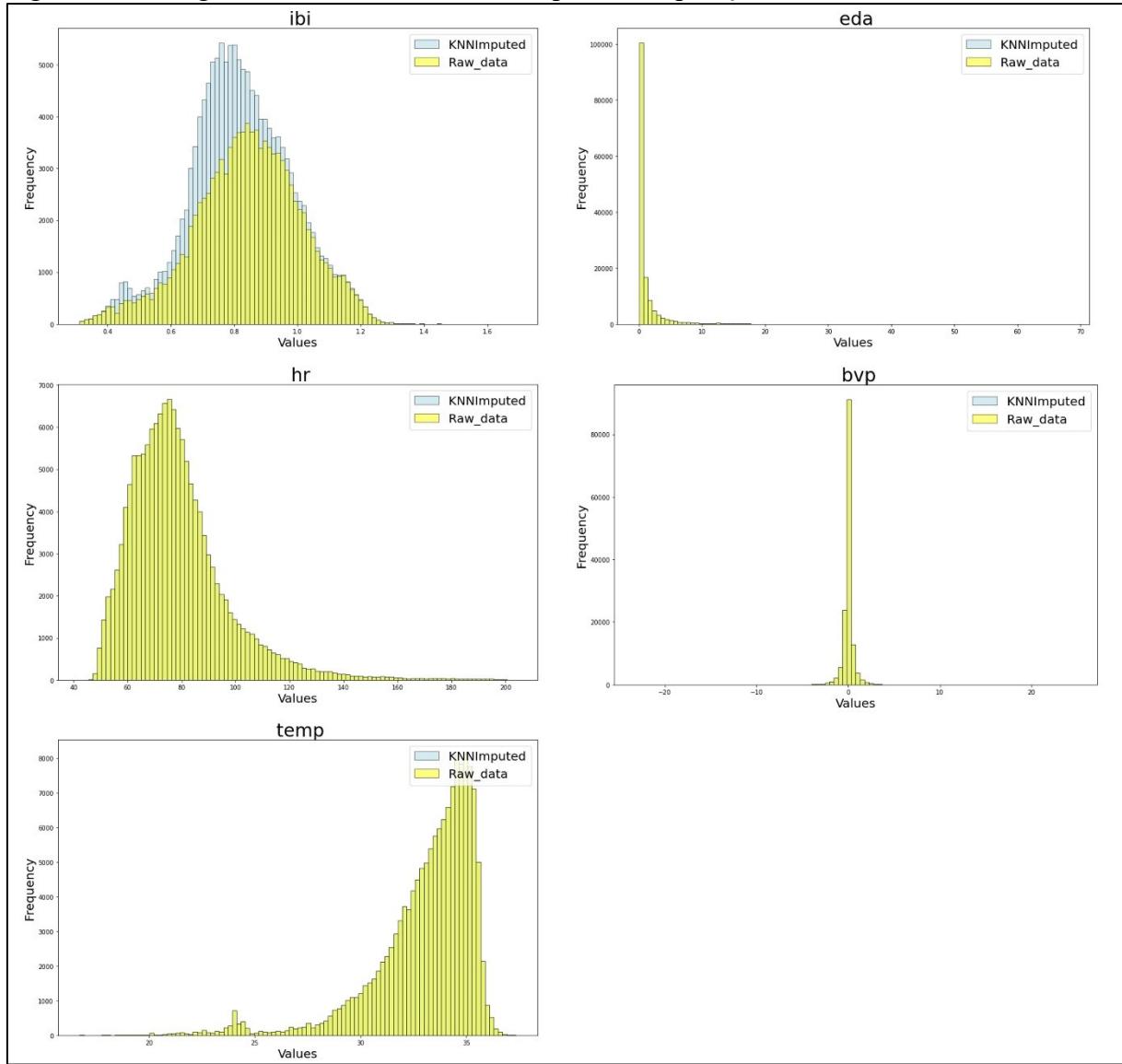
A base de dados resultante da transformação das leituras brutas dos sensores, que estavam disponíveis no PhysioNet, possui uma dimensão de 196.336 instâncias contendo 7 atributos, 6 independentes e 1 dependente, como descrito com maiores detalhes na seção de metodologia. Ao realizar-se a transformação de cada leitura bruta dos sensores para uma média de seus valores, numa janela de um minuto, seguida da junção dos dados dos diversos sensores de um mesmo indivíduo, observou-se que em alguns minutos da série temporal, não havia quaisquer leituras. Nesses casos, as instâncias foram excluídas da base de dados, resultando em uma base com 144.535 instâncias.

Uma análise adicional, entretanto, mostrou que há casos em que ocorrem dados faltantes em um ou dois atributos de uma determinada instância. No entanto essa ocorrência não é uniforme, a quantidade de dados faltantes variou muito entre os atributos da base, enquanto os atributos EDA e BVP não possuíam dados faltantes e os atributos HR e temperatura possuíam 23 e 2 valores faltantes respectivamente, o atributo IBI possuía 35.013 valores faltantes, pouco mais de 24% dos valores. Para todos esses casos, novos dados foram imputados através do método dos vizinhos mais próximos. A distribuição dos dados antes e após a imputação está mostrada na Figura 7. Como pode ser observado, a imputação causou mínimos impactos na distribuição dos dados da maior parte das variáveis, exceto da variável IBI cuja média deslocou-se um pouco mais para esquerda da distribuição. A Figura 7 também mostra aspectos interessantes do perfil dos dados — a maior parte dos dados não apresenta uma distribuição normal. Apenas a variável IBI parece ter uma distribuição que se assemelha a uma distribuição normal. A variável BVP possui uma distribuição em forma de uma curva leptocúrtica, enquanto as variáveis TEMP e HR são assimétricas, TEMP com assimetria negativa e HR com assimetria positiva. Já a variável EDA apresenta um conjunto de valores com uma distribuição muito similar a uma distribuição exponencial.

Uma outra característica dessa base de dados é a ausência de valores duplicados, evitando a necessidade de qualquer tratamento adicional. Observamos também que não há correlações fortes entre a maior parte dos atributos da base (Fig. 8A). Enquanto isso há atributos que apresentam pouca ou nenhuma correlação com os outros atributos. O atributo BVP, por

exemplo, apresenta uma correlação quase nula com todos os outros atributos da base. Também podemos observar que a base de dados possui uma distribuição de dados bastante homogenia em relação às classes (Figura 8B, diagonal direta).

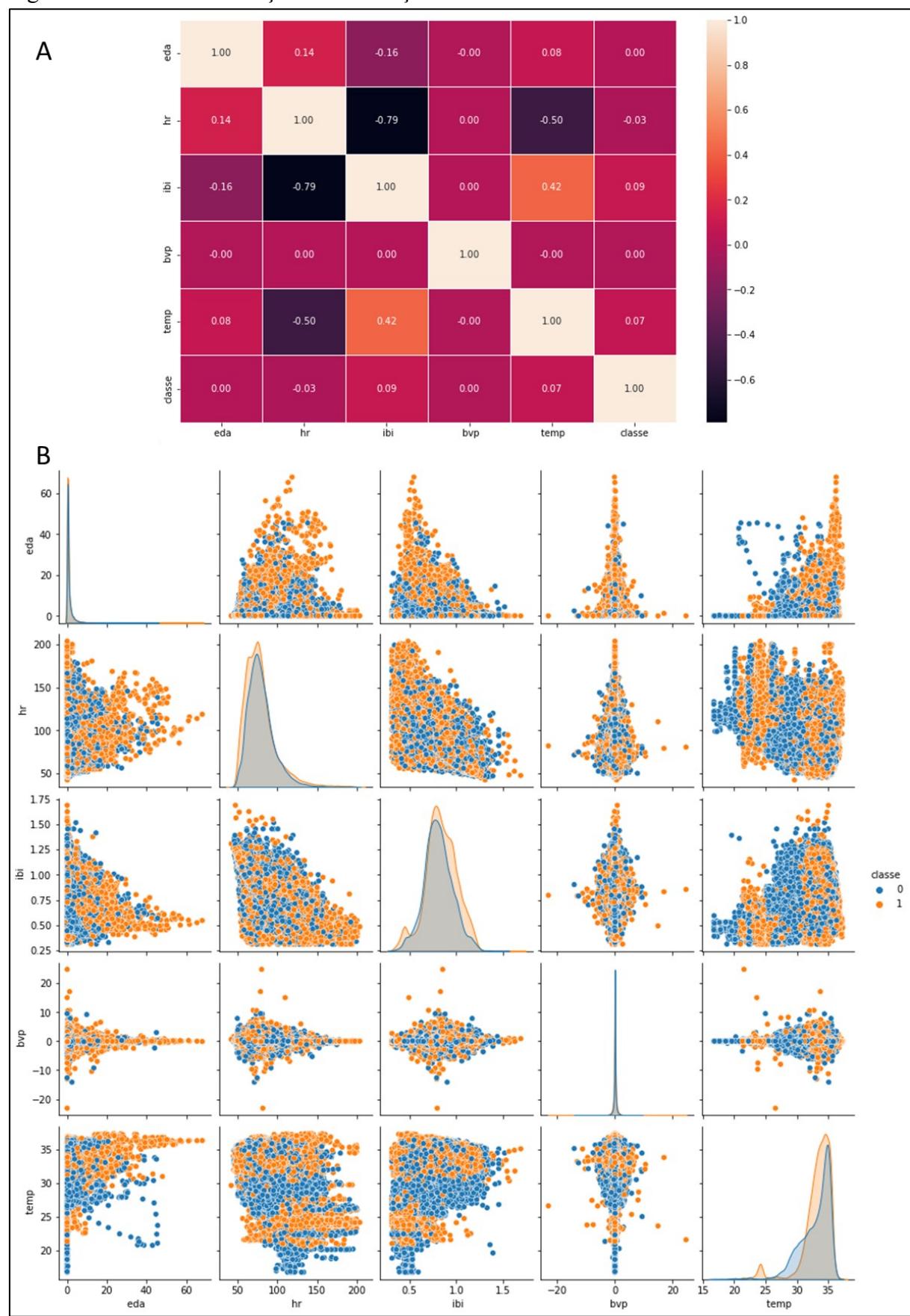
Figura 7 – Histograma dos atributos antes e depois da imputação de dados.



Legenda: No eixo X estão mostrados os valores dos atributos para cada instância, no eixo Y é mostrada a quantidade de instâncias que possuem uma determinada faixa de valor. Podemos ver que a imputação teve maior impacto na variável IBI – *Interbeat Interval*.

Antes de iniciar as etapas de treinamento e teste dos algoritmos, foi realizada uma otimização dos hiperparâmetros. As melhores configurações para cada algoritmo estão mostradas na Figura 9.

Figura 8 – Perfil de correlação e distribuição dos atributos da base de dados.



Fonte: Elaborado pelo autor

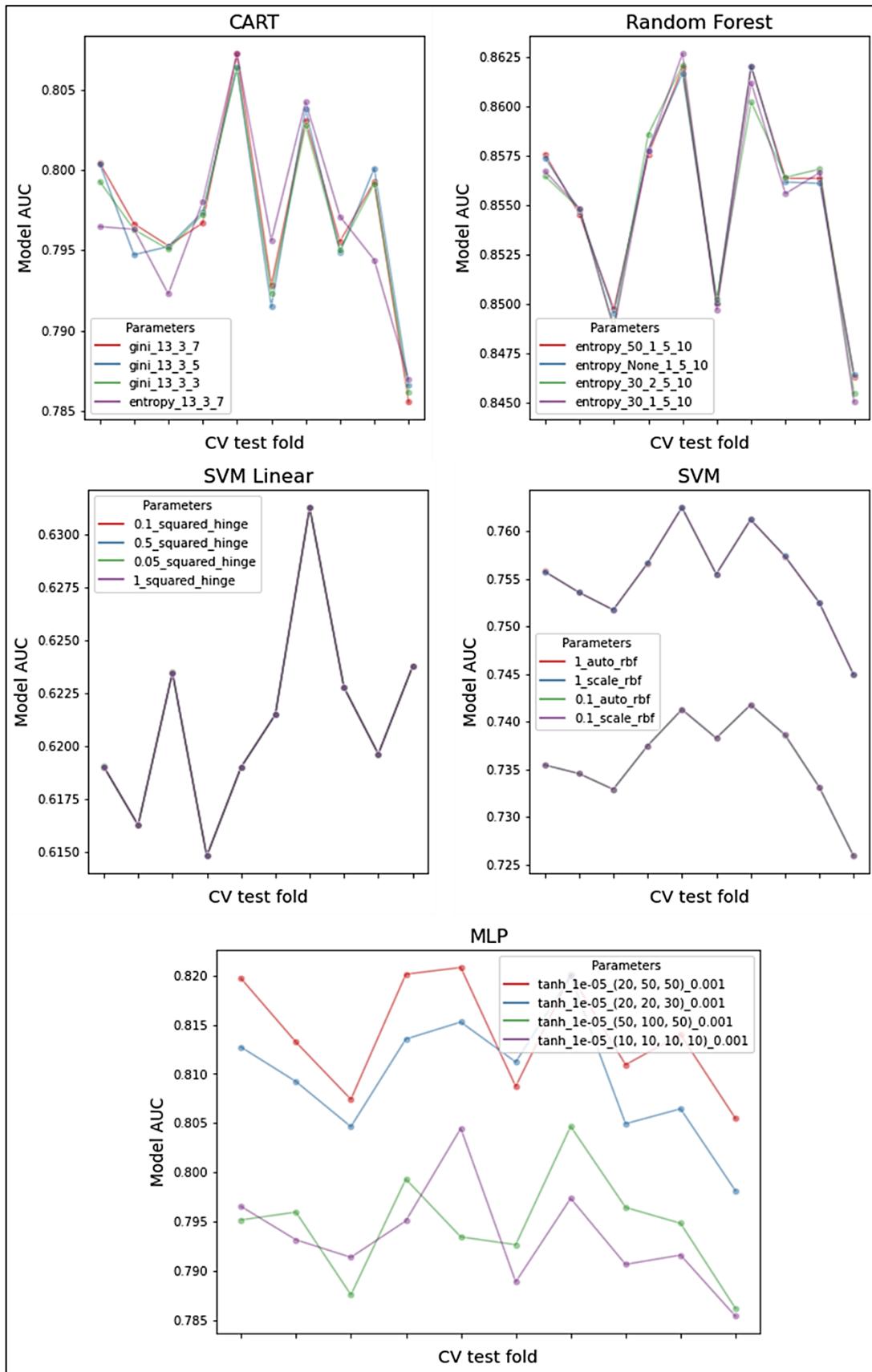
Para o algoritmo CART, implementado pela classe `DecisionTreeClassifier`, os melhores valores foram: `criterion = gini`, o que especifica o *Gini impurity*, que indica a chance de novos dados serem classificados de forma incorreta caso lhes fosse dado um valor aleatório de classe de acordo com a distribuição de classe observada; `min_samples_split = 7`, que especifica o menor número de amostras necessárias para que um nó interno seja dividido; `min_samples_leaf = 3`, que especifica o número mínimo de amostras que podem estar em um nó folha; `max_depth=13`, o que quer dizer que a árvore não será expandida até o máximo possível, ela só poderia ser expandida até 13 nós.

Para o comitê de árvores de decisão, implementado pela classe `RandomForestClassifier`, os melhores valores foram `criterion = entropy`, uma medida que irá indicar a desordem de um determinado nó para medir a qualidade da divisão daquele nó; `max_depth=50`, o que quer dizer que cada árvore na floresta aleatória não será expandida até o máximo possível, ela só poderia ser expandida até 50 nós; `min_samples_split = 5`, que especifica o menor número de amostras necessárias para que um nó interno seja dividido; `min_samples_leaf = 1`; que especifica o número mínimo de amostras que podem estar em um nó folha, neste caso, apenas uma amostra.

Os hiperparâmetros variam bastante entre os dois tipos de algoritmos de máquinas de vetores de suporte testados. A variação dos hiperparâmetros do `LinearSVC` praticamente não afetou a predição do algoritmo, o que pode ser visto por uma série de curvas sobrepostas (painele central à esquerda). O contrário foi observado com os hiperparâmetros do algoritmo `SVC`, onde os melhores valores foram, `kernel=rbf`; que especifica a função de decisão do algoritmo e permite a solução de problemas mais complexos; `C=1`; que é um parâmetro de regularização, quanto maior mais estreita é a margem de decisão da função, quanto menor, mais simples a função de decisão; `gamma = auto`, que define o quão longe irá a influência de um exemplo de treinamento, ou seja, de qual será o tamanho da influência das instâncias selecionadas como os vetores de suporte. O valor de `gamma` igual a “auto” significa que `gamma` assume o resultado da divisão: $1/\text{número_de_atributos}$. Logo, neste caso, `gamma` é igual a 0,2.

Em relação às redes neurais, ficou claro que um maior número de camadas ocultas reduziu a qualidade do modelo, assim como o uso de mais de 50 neurônios por camada. O melhor resultado foi obtido com o uso de 3 camadas ocultas, tendo 20 neurônios na primeira camada oculta, 50 na segunda e 50 na terceira. A melhor função de ativação (parâmetro que decide se o neurônio vai ser ativado ou não) dada pelo hiperparâmetro “`activation`” foi a função hiperbólica tangencial, representada pelo valor “`tanh`”. A melhor configuração do modelo também conta com uma taxa de aprendizado (`learning_rate_init`) de 0,001 e um valor `alpha = 1x10^-5`, que especifica um termo de regularização menor que o padrão para esse algoritmo.

Figura 9 – Resultado da otimização dos hiperparâmetros dos modelos testados.

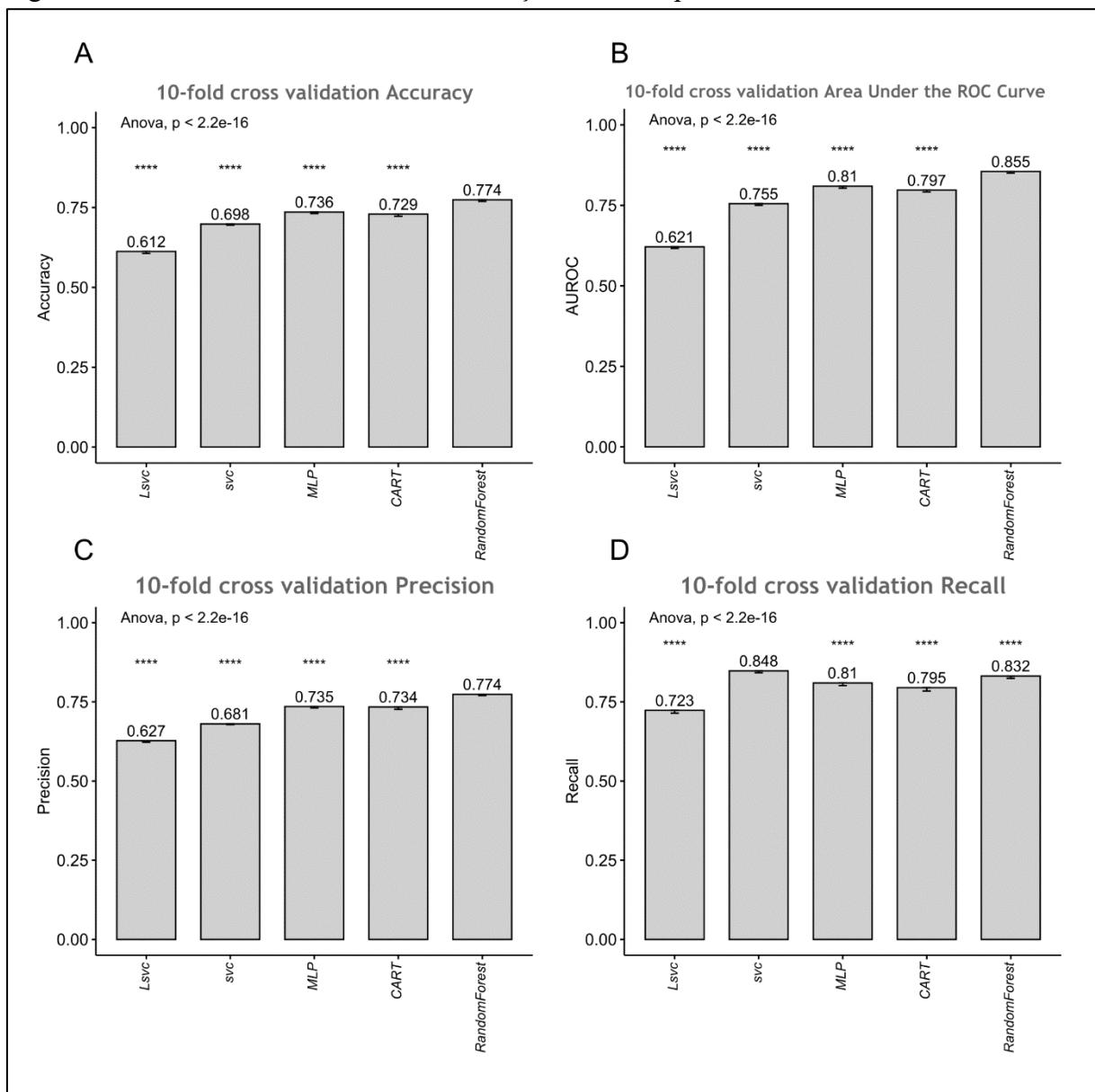


Fonte: elaborado pelo autor

5.2 ANÁLISE DA PROVA CONCEITUAL E DA QUALIDADE DOS MODELOS PARA CLASSIFICAÇÃO DA PRÉ-DIABETES

Os modelos, já com seus hiperparâmetros otimizados, foram submetidos a validação cruzada dentro do conjunto de treinamento e, para cada modelo, 4 métricas foram extraídas. Os resultados, mostrados na Figura 10, indicam que é possível utilizar aprendizagem de máquina e dados de dispositivos vestíveis para classificar indivíduos portadores da pré-diabetes e indivíduos saudáveis. Apesar disso, essa classificação não é perfeita, a acurácia da predição não é elevada. Ela variou entre 0,612 até 0,774, o que quer dizer que no melhor caso, 77,4% das previsões foram corretas e 22,6 % foram incorretas.

Figura 10 – Resultado das métricas de avaliação de desempenho dos modelos.



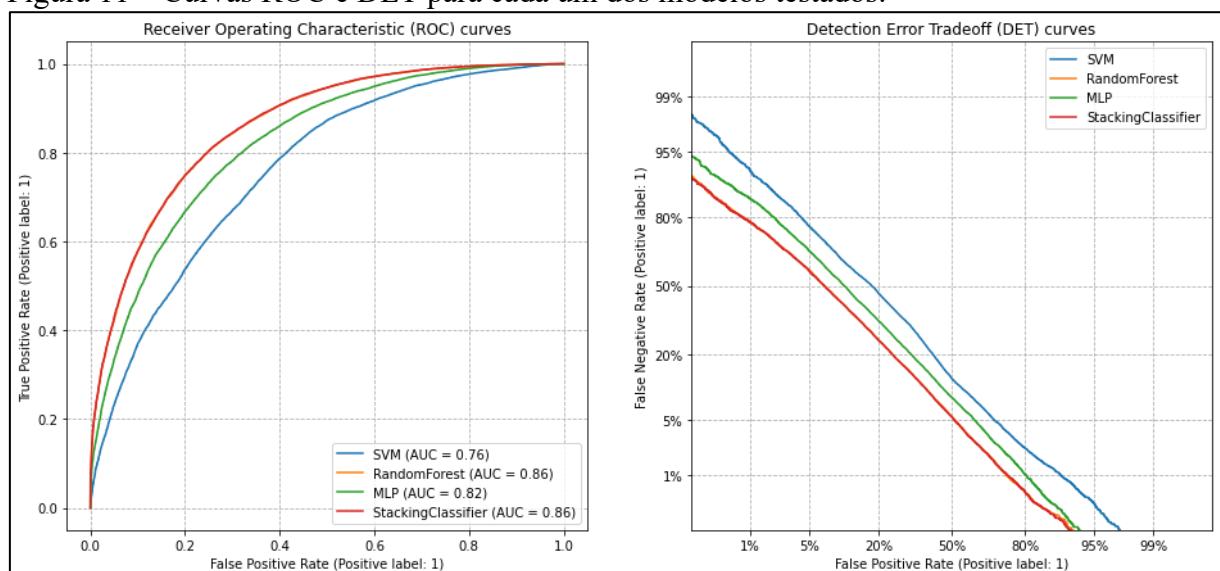
Fonte: elaborado pelo autor

Legenda: Os asteriscos representam a significância estatística da comparação entre duas médias. Neste caso, cada valor está comparado com o valor do algoritmo com o maior score em cada uma das 4 métricas. *** representa um $p-value < 0,0001$.

O algoritmo *Random Forest* teve uma performance significativamente melhor que os outros modelos testados em termos de acurácia (Fig. 10A) e precisão (Fig. 10C), ou seja, o modelo possui uma melhor capacidade de não classificar como positivo, algo que é negativo. Isso quer dizer que o modelo tem a menor taxa de falso positivo entre todos os modelos. Em outras palavras ele é o algoritmo que menos classifica dados de pacientes sem pré-diabetes como sendo dados de pré-diabéticos. Apesar disso, o *Random Forest* teve uma sensibilidade um pouco menor que o algoritmo SVC (Fig. 10 D). Além de uma maior acurácia e precisão o *Random Forest* também tem o melhor equilíbrio entre a taxa de falsos positivos e a taxa de positivos verdadeiros, representado por uma AUC (*area under the curve*) de 0,855, significativamente maior que a dos outros algoritmos (Fig. 10B).

Uma vez que o Random Forest possui a melhor acurácia, precisão e AUC dos 3 algoritmos, mas possui uma sensibilidade inferior ao SVC, tentou-se criar um classificador misto que utilizava as entradas do SVC, do Random Forest e da rede neural MLP. O gráfico da curva ROC para esses três métodos, assim como o do classificador heterogêneo está mostrado na Figura 11. Fica claro que a criação do comitê heterogêneo não gerou nenhuma melhora em relação a utilização única do algoritmo Random Forest, mas gerou aumento do custo computacional, o que nos leva a descartar esse modelo. A curva DET (Figura 11) deixa claro que há uma sobreposição entre a performance do classificador heterogêneo e do Random Forest, havendo diferença apenas em valores muito baixos de falsos positivos e negativos.

Figura 11 – Curvas ROC e DET para cada um dos modelos testados.

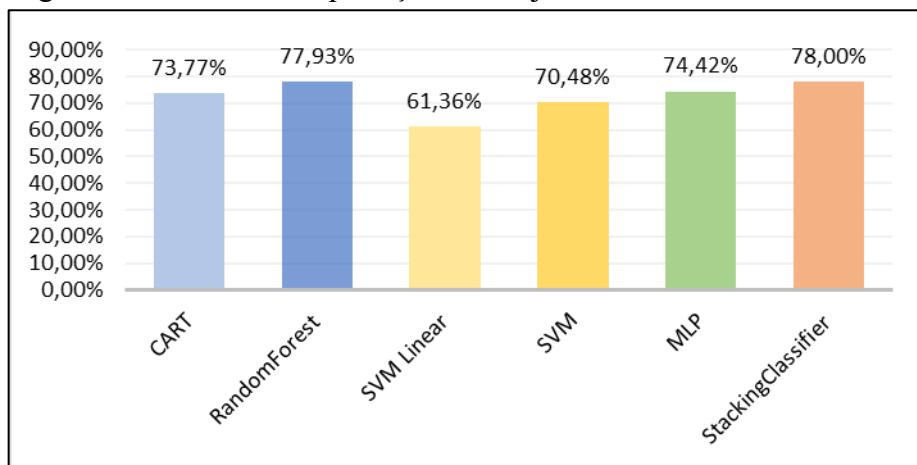


Fonte: Elaborado pelo autor

Nota: A curva ROC mostra o balanço entre positivos verdadeiros e falso positivos, a área sob a curva ROC é um indicador da qualidade do modelo, quanto maior a área melhor o modelo. A curva DET mostra o comportamento do modelo de acordo com a taxa de falsos positivos contra a taxa de falsos negativos, fica claro que para obter-se uma baixíssima taxa de falsos negativos, ou seja, dados indicativos de pré-diabetes classificados como saudáveis, há uma alta taxa de falsos positivos. Observa-se no gráfico que as linhas de *Stacking Classifier* e do *Random Forest* estão sobrepostas.

Após a execução da validação cruzada no conjunto de treino, os modelos também foram comparados em relação ao conjunto de teste. Como esperado, o algoritmo de Random Forest teve uma melhor performance em relação aos outros algoritmos testados. A comparação com um comitê heterogêneo de classificação também mostrou a efetividade do Random Forest, pois a diferença de acurácia no seu uso, sozinho ou em conjunto com outros 3 algoritmos, foi de apenas 0,07%.

Figura 12 – Acurácia da predição no conjunto de teste.



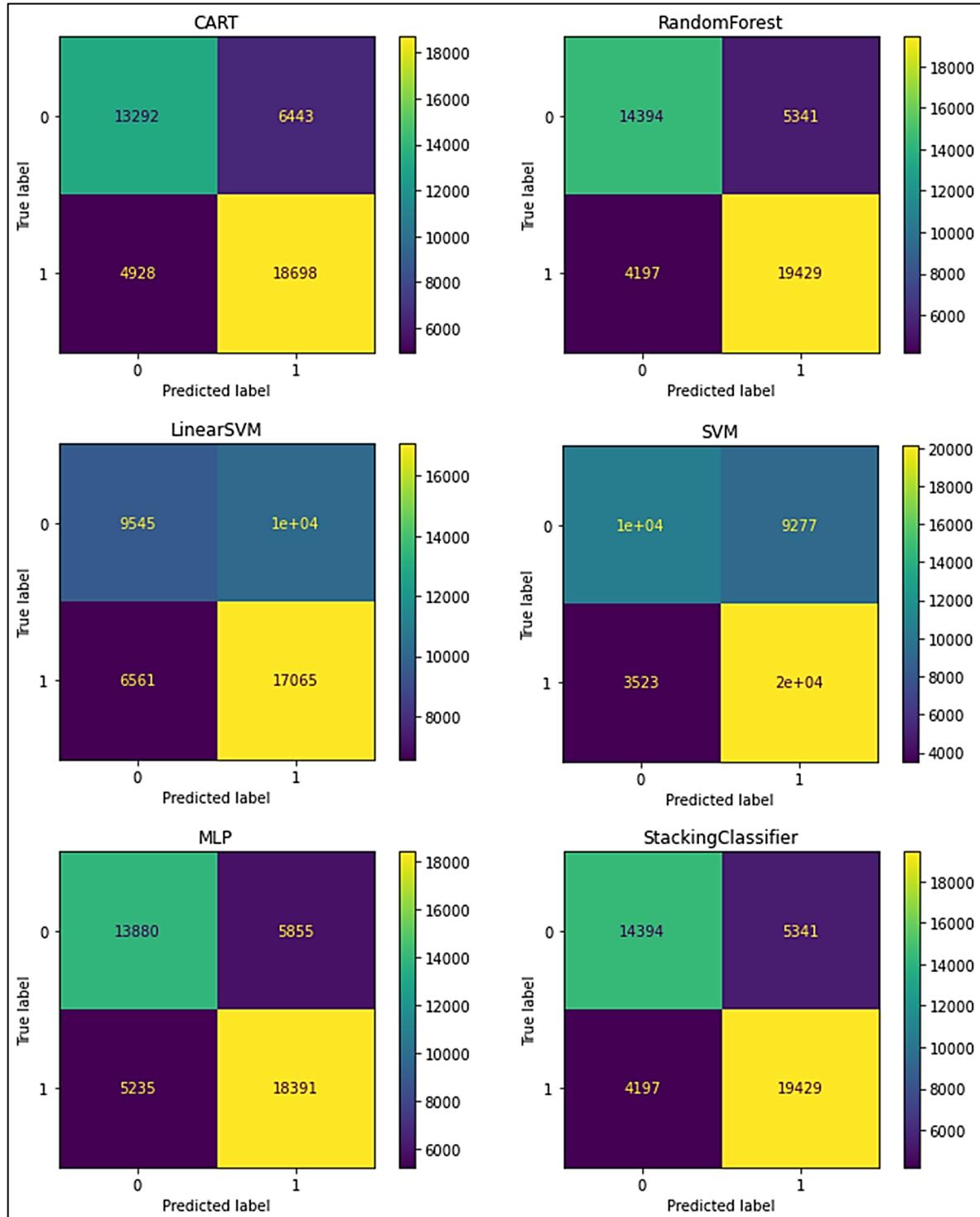
Fonte: Elaborado pelo autor

Os resultados detalhados das predições de cada um dos algoritmos avaliados podem ser vistos na Figura 13, que mostra as matrizes de confusão geradas. Fica claro que o algoritmo SVC produz um número muito menor de erros em relação a dados de pré-diabéticos (rótulo 1) classificados como dados de indivíduos saudáveis (rótulo 0), apenas 3523 erros, num universo de 23626 dados de pré-diabéticos, uma taxa de falso negativo de 14,91%. Apesar disso, ele apresenta uma das maiores taxa de falso positivo dentre todos os algoritmos, 46,96%, um valor muito maior que o apresentado pelo Random Forest cuja taxa é de 27,06%. Isso reflete uma baixa precisão (68,42%) no algoritmo SVC com o kernel RBF.

Como para um teste de diagnóstico deve haver um equilíbrio entre o número de falsos positivos e falsos negativos e, uma vez que o SVC produz uma taxa de falso positivo muito grande (quase um erro a cada duas estimativas), apesar da sua baixa taxa de falso negativo, ele é menos adequado para uma tarefa de predição da pré-diabetes que o algoritmo *Random Forest*.

que apresentou a maior área sob a curva ROC nos dados de treino (Figura 11) e apresenta uma taxa de falso positivo de 27,06% e uma taxa de falso negativo de 17,76%.

Figura 13 – Matrizes de confusão resultantes da predição do conjunto de teste.



Fonte: Elaborado pelo Autor

Legenda: 0 representa a classe “saudável” e 1 representa a classe “pré-diabético”

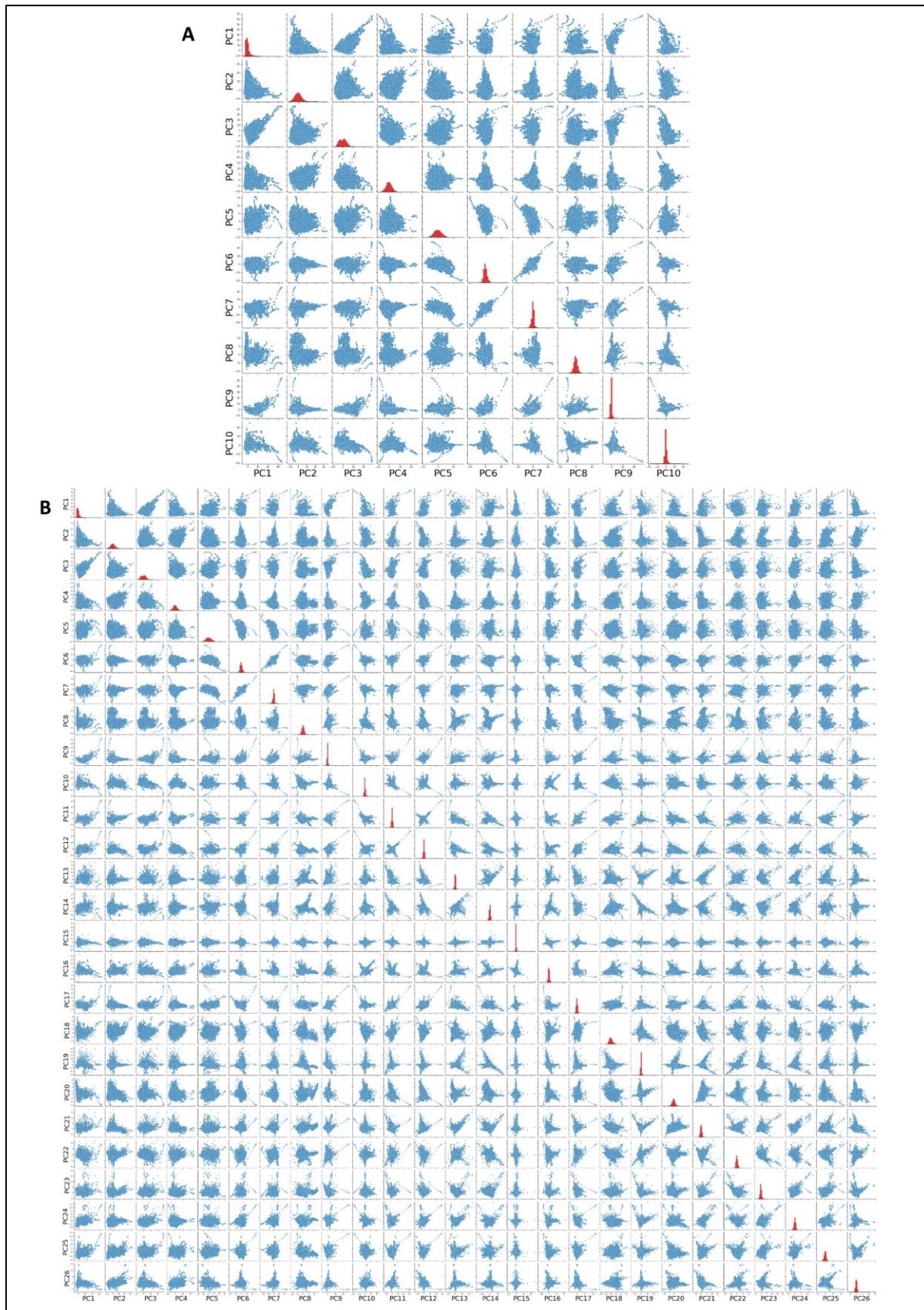
5.3 ANÁLISE DA BASE CONTENDO OS DADOS DECOMPOSTOS PELO PACOTE FLIRT E APLICAÇÃO DO ALGORITMO DE RANDOM FOREST

A base de dados resultante do uso do FLIRT para decomposição de sinais possui uma estrutura bem distinta da base de dados das leituras brutas dos sensores, cujos dados estavam disponíveis no PhysioNet. Enquanto esta possuía apenas 7 atributos, sendo 6 independentes e 1 dependente, a base contendo os sinais decompostos possui 99 atributos, sendo 98 independentes e 1 dependente. Os atributos dessa nova base de dados são representados apenas por variáveis numéricas e pela variável categórica binária ‘sexo’. Em relação a quantidade de instâncias, enquanto a base de dados brutos possuía 144.535 instâncias a base de dados do FLIRT possui apenas 86437. Essa redução de aproximadamente 40% do número de instâncias reflete a presença de ruídos, artefatos e a baixa qualidade de sinal dos sensores em determinados momentos da série temporal.

Devido a elevada quantidade de atributos, foi realizada a redução de dimensionalidade através da aplicação do algoritmo de PCA. Foram geradas duas bases de dados após a aplicação do PCA, na primeira delas foram mantidos apenas os 10 principais componentes, que respondiam por 80,48% da variância contida na amostra; na segunda foram mantidos um número de componentes que explica 95% da variância da amostra, valor esse que foi obtido com o uso de 26 componentes principais, ou seja, pouco mais de um quarto do total de atributos presente na base de dados processada pelo FLIRT. A Figura 14 mostra a distribuição dos novos atributos, percebe-se que boa parte deles apresenta uma distribuição normal, embora haja, em alguns, a presença do fenômeno de curtose. Também fica claro que os atributos não são correlacionados, ou possuem apenas uma fraca correlação, evidenciando o efeito do PCA.

Foi realizada uma otimização dos hiperparâmetros do algoritmo derivado do melhor modelo obtido a partir da base de dados dos sinais brutos, o algoritmo de *Random Forest*. Para base completa e para ambas as bases de dados, com 10 e 26 componentes principais, os melhores valores de parâmetros foram `criterion = entropy`, `min_samples_split = 2` e `min_samples_leaf = 1`. Quanto a profundidade máxima da árvore, não houve diferença de score para valores acima de 50 para as bases reduzidas, como pode ser observado na Figura 15. Nota-se nesta figura, que as linhas vermelha e azul (que representam a profundidade de 50 e ausência de poda por profundidade (valor “None”)) estão totalmente sobrepostas. Apesar disso, uma árvore podada é mais simples e por isso, o valor de 50 foi escolhido para o hiperparâmetro `max_depth`. O valor de 30 foi escolhido para `max_depth` na base de dados completa.

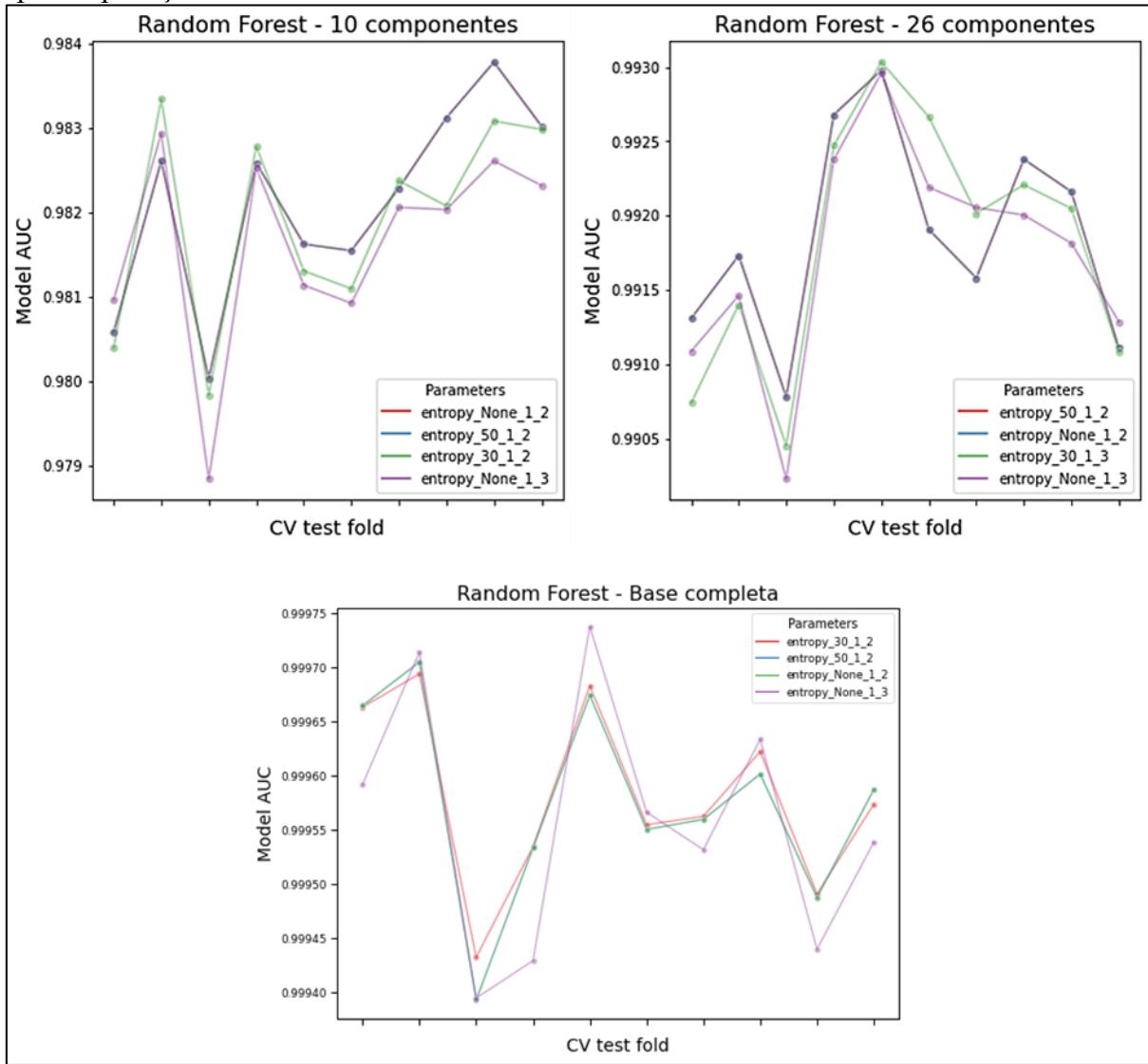
Figura 14 - Correlação e distribuição dos atributos das bases após a aplicação do FLIRT e PCA.



Fonte: Elaborado pelo Autor

Legenda: Os quadros da linha diagonal em vermelho mostram um histograma de cada atributo, o restante dos quadros mostra comparações par a par de variáveis. Base com 10 (A) e 26 (B) componentes principais.

Figura 15 - Resultado da otimização dos hiperparâmetros do Random Forest nas bases de dados após a aplicação do FLIRT e PCA.

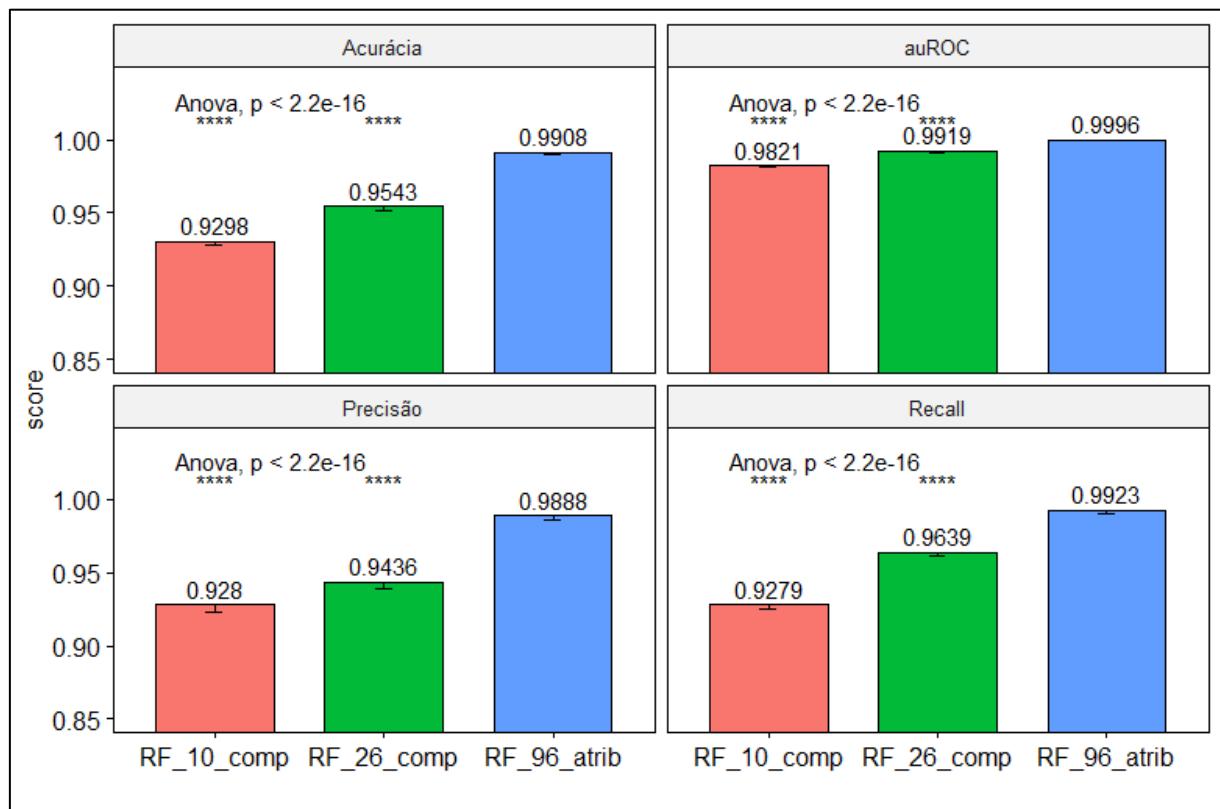


Fonte: Elaborado pelo autor

Os modelos, já com seus hiperparâmetros otimizados, foram submetidos a validação cruzada dentro do conjunto de treinamento das duas bases de dados. Assim como na análise com os dados brutos dos sensores, nessas novas bases também foram extraídas quatro métricas de performance dos modelos. Os resultados, mostrados na Figura 16, indicam uma superioridade no modelo construído sobre a base de dados contendo os 26 componentes principais. Para todas as quatro métricas avaliadas, há diferença estatisticamente significativa entre a base contendo 10 e a base contendo 26 atributos. Quando comparamos esses resultados com os resultados da validação cruzada utilizando o modelo Random Forest sobre a base de dados dos sinais brutos, fica claro que a utilização do FLIRT para decomposição de sinal da dos sensores aumenta em mais de 20% a acurácia de predição do modelo, chegando a valores

superiores a 90%. A utilização de uma base com 26 componentes principais em conjunto com o modelo de Random Forest teve uma sensibilidade de 96,4% na detecção de pré-diabetes, ou seja, apenas 3,6% dos padrões fisiológicos indicativos de pré-diabetes foram classificados erroneamente. O que indica que a aplicação do FLIRT sobre os dados de IBI, EDA foi capaz de gerar bons biomarcadores digitais para a doença. Isso deixa claro que é possível utilizar aprendizagem de máquina e dados processados de dispositivos vestíveis para classificar indivíduos portadores da pré-diabetes e assim alertar quando determinados parâmetros fisiológicos de uma pessoa estão indicando a entrada numa fase de risco para o desenvolvimento do diabetes mellitus tipo 2.

Figura 16 - Resultado da validação cruzada com o algoritmo de Random Forest nas novas bases de dados.



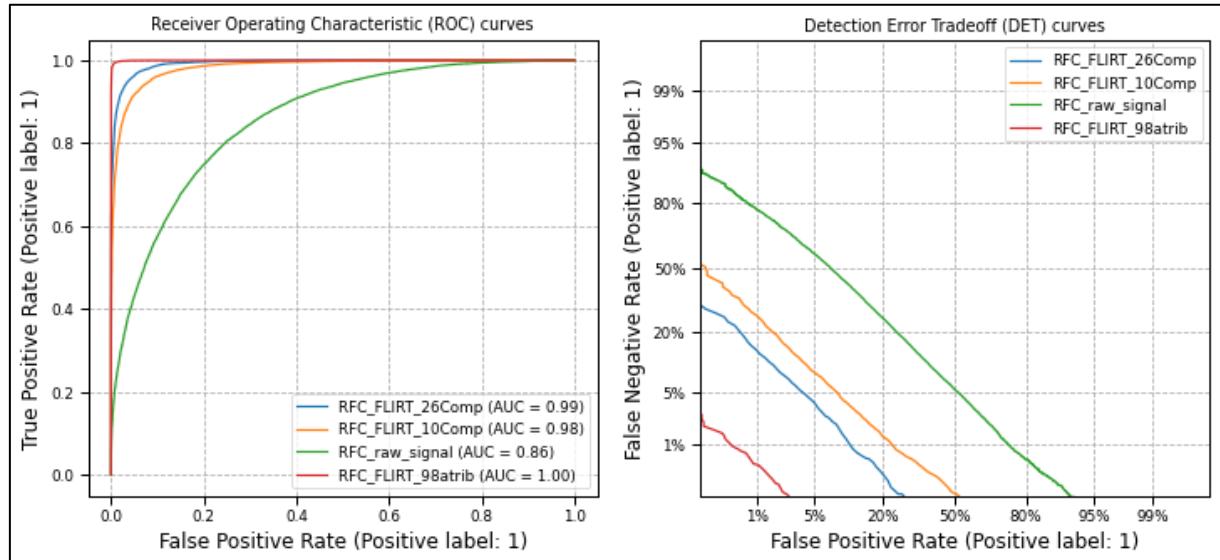
Fonte: Elaborado pelo Autor

Legenda: “RF_10_comp” refere-se à aplicação do Random Forest sobre a base de dados onde foram mantidos os 10 principais componentes, já o “RF_10_comp” refere-se à base onde foram mantidos 26 componentes principais. O “RF_96_atrib” refere-se a base completa. O *p-value* mostrado refere-se à aplicação do teste de ANOVA e **** representa um *p-value* de 0,0001 no teste *t* de Student para comparação entre as médias em relação à média com maior valor em cada métrica.

A Figura 17 mostra como a aplicação do FLIRT alterou a área sob a curva ROC em relação a utilização da base de dados com sinais brutos. A curva DET também deixa claro como a utilização de 26 componentes é bem melhor que a utilização de 10 componentes principais, e

como a utilização da base completa supera em eficácia a utilização das bases com dimensionalidade reduzida.

Figura 17 – Curvas ROC e DET construída a partir do uso de Random Forest em bases com diferentes dimensionalidades.



Fonte: Elaborado pelo Autor

Legenda: O padrão de nomes “RFC_FLIRT...” refere-se aos modelos aplicados base de dados cujos sinais de IBI e EDA foram processados pelo pacote FLIRT, já o modelo “RFC_raw_signal” refere-se à aplicação do Random Forest na base de dados com os sinais brutos captados pelos sensores da Pulseira Empatica E4.

A aplicação do modelo treinado no conjunto de testes gerou scores igualmente altos em relação a validação cruzada realizada no conjunto de treinamento, eliminando a possibilidade da ocorrência de um super ajustes do modelo aos dados do conjunto de treinamento. A acurácia da predição do conjunto de teste foi de 93,38% utilizando a base de dados com 10 componentes principais. Já com a base contendo os 26 atributos, a acurácia da predição foi de 95,77%, sendo necessário apenas 50% a mais de tempo para completar a tarefa de treino em relação a base menor. A utilização da base completa contendo 98 atributos, por outro lado, levou 68% a mais de tempo que a base contendo apenas 10 componentes, gerando uma acurácia de 99,3%. Já a tarefa de predição levou aproximadamente o mesmo tempo para os três modelos. A Tabela 1 contém todas as 4 métricas previamente avaliadas, agora aplicas no conjunto de teste.

Tabela 1 – Métricas da aplicação do Random Forest para predição no conjunto de teste

	Acurácia	Precisão	Recall	AUROC
Base com 10 componentes	93,38 %	93,11%	93,26%	98,31%
Base com 26 componentes	95,77%	94,75%	96,65%	99,29%
Base com 98 atributos	99,3%	99,13%	99,42%	99,97%

Fonte: Elaborado pelo Autor

Este estudo prova que é possível utilizar ambos os dados de pulseiras inteligentes, brutos e processados, para a tarefa de diferenciação de pessoas saudáveis e pessoas portadoras de pré-diabetes utilizando o algoritmo de Random Forest. Apesar disso a utilização de dados processados pelo FLIRT aumentou muito a taxa predições corretas do algoritmo.

Essa tarefa de diferenciação foi previamente investigada por dois estudos. No primeiro, Prabha *et al.*, (2021a) obtiveram uma alta acurácia na predição do estado de saúde dos indivíduos, valores de 83,87% e 84,49% ao utilizar algoritmos e KNN e SVM. Esse valor, apesar de alto, é bem mais baixo que o encontrado no presente trabalho com o uso do FLIRT. Prabha *et al.*, (2021a) também utilizaram dados diferentes dos nossos. Apesar dos dados utilizados por esses autores derivarem do mesmo modelo de pulseira, a Empatica E4, foram utilizados apenas dados coletados a partir do sensor de fotopletismografia, não utilizando dados de atividade eletrodérmica ou de temperatura. Os dados de fotopletismografia também não foram utilizados no seu estado bruto, os autores aplicaram um processamento a partir dos MFCCs, um conjunto de características amplamente utilizadas em reconhecimento de fala e processamento de sinais de áudio. Embora tenham utilizado apenas o sinal de fotopletismografia, os autores também utilizaram outros atributos como peso, altura e idade dos indivíduos, características que não foram utilizadas por nós no atual estudo. Ao fim, Prabha *et al.*, (2021a) obtiveram uma base com 107 atributos. Esses autores também utilizaram a técnica de PCA para redução de dimensionalidade da base de dados, elevando a acurácia da classificação com SMV de 84,49% para 92,28%. Esse comportamento é inverso ao observado no presente estudo, onde a aplicação do PCA, embora tenha reduzido levemente o tempo de treinamento do modelo, também reduziu sua acurácia.

Outra diferença em relação ao nosso estudo, é que Prabha *et al.*, (2021a) utilizaram dados de pessoas saudáveis, diabéticas e pré-diabéticas visando um tarefa de classificação nesses três estados. Sua base, no entanto, estava bastante desbalanceada em relação a pessoas pré-diabéticas, que correspondiam a apenas 7,37% dos indivíduos, isso dificulta uma comparação direta com nosso modelo, pois não sabemos se o modelo apresentado pelos autores tem maior percentual de acerto em relação aos indivíduos saudáveis, diabéticos ou aos pré-diabéticos.

Posteriormente os mesmos autores publicaram um novo estudo, onde testaram métodos combinados para redução de dimensionalidade em conjunto com o algoritmo de XGBoost, obtendo uma acurácia de 99,93% na predição dos estados de saúde (PRABHA *et al.*, 2021b),.

um valor superior ao encontrado no presente trabalho. Essa alta acurácia foi obtida na predição de todas as classes: saudáveis, diabéticos e pré-diabéticos. Apesar de a fonte de dados ser a mesma pulseira, os dados adicionais utilizados pelos autores foram distintos, enquanto nós utilizamos o sexo, Prabha *et al.*, (2021b) utilizaram dados de peso, altura e idade, o que pode ter maximizado a acurácia de seu modelo. Outra diferença entre os trabalhos está na população avaliada, enquanto os autores utilizaram dados de 217 pacientes, nós conseguimos chegar a uma acurácia também superior a 99% com dados de apenas 15 pacientes. É possível que um número de pacientes analisados também eleve a acurácia do modelo de *Random Forest* obtido neste trabalho.

Uma das vantagens da existência de dois estudos, Prabha *et al.* (2021b) e o presente estudo, realizados com duas populações distintas e utilizando métodos distintos de processamento dos sinais brutos derivados de pulseiras inteligentes comprovar que esses sinais são bons preditores e podem servir de biomarcadores digitais para detecção de pré-diabetes em pacientes do mundo todo.

O presente estudo, no entanto, não está livre de limitações. Como foram utilizados dados públicos presentes em bases de dados, não tivemos a opção de escolha de indivíduos do estudo, assim como não pudemos realizar aumento do número de indivíduos amostrados nem recuperar leituras de diferentes períodos. Uma das limitações em relação a quantidade de dados reside no fato de serem utilizados dados de apenas 15 indivíduos, um valor que não abrange toda a variedade de tonalidade de cor de pele e etnias presente na população global, da mesma forma que anteriormente citado por (BENT *et al.*, 2021b). Isso, no entanto, não quer dizer que a captura de novos dados não será capaz de gerar modelos com performances similares frente a um maior número de indivíduos.

Outra limitação que pode impactar nas altas métricas obtidas neste estudo é o tipo de validação utilizada. Neste estudo foi utilizada uma validação cruzada de 10 rounds, com uma divisão de instâncias totalmente aleatória, além de um grupo de teste à parte. Embora cada instância da base represente um tempo de coleta do indivíduo, que reflete o estado fisiológico daquele indivíduo apenas num determinado instante, de forma que cada instância represente apenas uma “fotografia” da saúde do indivíduo naquele instante, é possível que o modelo se ajuste bastante ao padrão de dados fisiológicos das 15 pessoas avaliadas e que a performance seja parcialmente reduzida quando aplicado a dados de novos indivíduos. Ao escolher esse tipo de validação, foi levado em conta que outros estudos publicados também utilizaram validação

cruzada de 10 rounds, como o estudo realizado por Prabha *et al.* (2021b). Além disso, levamos em consideração que uma redução de uma fração de pessoas, para atuar como um conjunto de validação de dados, pode reduzir muito a variabilidade de dados que teríamos a nossa disposição para treinamento dos modelos. No entanto, consideramos esta limitação como um elemento que poderia introduzir um certo viés no estudo. Uma alternativa para eliminação desse viés seria a fusão de dados de mais de um banco que utilizasse os mesmos critérios e equipamentos de coleta de dados, o que aumentaria o número de indivíduos disponíveis. Outra alternativa seria realizar uma validação do tipo *Leave-one-person-out cross validation* (LOOCV), onde cada pessoa como atuaria como um conjunto de validação em cada um dos rounds da validação cruzada. Esse método de validação vem se tornando uma alternativa em estudos com dados captados por sensores e derivados de estudos com seres humanos, onde nem sempre se pode aumentar o número de indivíduos utilizados nos experimentos (GHOLAMIANGONABADI; KISELOV; GROLINGER, 2020).

6 CONCLUSÃO

A busca por métodos não invasivos que sejam capazes de gerar alguma informação relevantes sobre o estado de saúde das pessoas vem se intensificando ano após ano. Com o avanço e popularização dos métodos de aprendizagem de máquina a descoberta de biomarcadores de doenças saiu exclusivamente do ambiente dos laboratórios de saúde e passou a fazer parte do campo de pesquisa da tecnologia da informação. A grande quantidade de dados sobre saúde, fez surgir um campo imenso de busca por biomarcadores digitais. Atualmente, esta é área ativa de busca no campo de prevenção e acompanhamento da diabetes.

Neste trabalho mostramos que é possível utilizar dados de atividade eletrodérmica, temperatura, sexo, frequência cardíaca, intervalo interbatidas e volume sanguíneo no pulso, derivados de sensores presentes na pulseira Empatica E4, para predição do pré-diabetes. Apesar disso, ficou claro que o uso de dados brutos para treinar modelos de aprendizado de máquina não gera bons classificadores, que embora sejam melhores que um classificador aleatório, ainda possuem altas taxas de erros. Por outro lado, o trabalho evidenciou que a aplicação da biblioteca FLIRT para processamento de dados brutos, seguido da aplicação do algoritmo de *Random Forest* é capaz gerar um ótimo balanço entre sensibilidade e especificidade. Como opção para redução de custo computacional também é possível reduzir a dimensionalidade de base mantendo os 26 componentes principais para classificar com alta confiança a instalação da pré-diabetes em um indivíduo a partir de dados de dispositivos vestíveis.

Apesar de atuar como prova de conceito, esse trabalho não fornece uma solução final para o problema da predição da pré-diabetes. Estudos futuros devem buscar aumentar o número de pessoas incluídas no estudo, assim como buscar maximizar a diversidade de marcas de relógios inteligentes incluídas no mesmo, a fim de tornar a solução mais democrática para a população em geral. Além de incrementar o número de amostras e variedades de dispositivos vestíveis, estudos futuros também devem estar atentos para a realização de uma validação cruzada do tipo LOOCV, não realizada neste estudo, além do método de validação por *k-fold cross validation*, a fim de reduzir possíveis vieses do estudo derivados do sobreajuste dos modelos de aprendizagem de máquina ao perfil fisiológico dos pacientes.

Por fim, apesar dos desafios em estudos amplos com variados participantes e dispositivos, é inegável a urgência desse tipo de pesquisa. Especialmente porque o estudo atual confirma a relevância desses parâmetros fisiológicos em atuarem como biomarcadores digitais valiosos para a detecção antecipada do desenvolvimento de uma condição cada vez mais prevalente na população, o pré-diabetes.

REFERÊNCIAS

AHMED, A. *et al.* Performance of artificial intelligence models in estimating blood glucose level among diabetic patients using non-invasive wearable device data. **Computer Methods and Programs in Biomedicine Update**, v. 3, n. January, p. 100094, 2023.

AMERICAN DIABETES ASSOCIATION. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2021. **Diabetes Care**, v. 44, n. Supplement_1, p. S15–S33, 1 jan. 2021.

BARBER, S. R. *et al.* Risk assessment tools for detecting those with pre-diabetes: A systematic review. **Diabetes Research and Clinical Practice**, v. 105, n. 1, p. 1–13, 2014.

BENT, B. *et al.* Engineering digital biomarkers of interstitial glucose from noninvasive smartwatches. **npj Digital Medicine**, v. 4, n. 1, p. 1–11, 2021a.

BENT, B. *et al.* Non-invasive wearables for remote monitoring of HbA1c and glucose variability: proof of concept. **BMJ Open Diabetes Research & Care**, v. 9, n. 1, p. e002027, 15 jun. 2021b.

BERG, J. M. *et al.* **Biochemistry**. 9^a ed. New York: Macmillan Learning, 2019.

BEULENS, J. *et al.* Risk and management of pre-diabetes. **European Journal of Preventive Cardiology**, v. 26, n. 2_suppl, p. 47–54, 26 dez. 2019.

BISWAS, D. *et al.* Heart Rate Estimation From Wrist-Worn Photoplethysmography: A Review. **IEEE Sensors Journal**, v. 19, n. 16, p. 6560–6570, 15 ago. 2019.

BOSCARI, F. *et al.* FreeStyle Libre and Dexcom G4 Platinum sensors: Accuracy comparisons during two weeks of home use and use during experimentally induced glucose excursions. **Nutrition, Metabolism and Cardiovascular Diseases**, v. 28, n. 2, p. 180–186, fev. 2018.

CHA, K. H. *et al.* Evaluation of Commercial Glucometer Test Strips for Potential Measurement of Glucose in Tears. **Analytical Chemistry**, v. 86, n. 3, p. 1902–1908, 4 fev. 2014.

CHAKI, J. *et al.* Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review. **Journal of King Saud University - Computer and Information Sciences**, v. 34, n. 6, p. 3204–3225, jun. 2022.

CHO, P. *et al.* **BIG IDEAs Lab Glycemic Variability and Wearable Device Data (version 1.1.0)**. Disponível em: <<https://physionet.org/content/big-ideas-glycemic-wearable/1.1.0/>>. Acesso em: 3 abr. 2023.

CORLISS, J. **Want to check your heart rate? Here's how**. Disponível em: <<https://www.health.harvard.edu/heart-health/want-to-check-your-heart-rate-heres-how>>. Acesso em: 26 maio. 2023.

DATYE, K. A. *et al.* Advances, Challenges, and Cost Associated with Continuous Glucose Monitor Use in Adolescents and Young Adults with Type 1 Diabetes. **Current Diabetes**

Reports, v. 21, n. 7, p. 22, 15 jul. 2021.

DEISS, D. *et al.* Clinical Practice Recommendations on the Routine Use of Eversense, the First Long-Term Implantable Continuous Glucose Monitoring System. **Diabetes Technology & Therapeutics**, v. 21, n. 5, p. 254–264, maio 2019.

DUAN, D.; KENGNE, A. P.; ECHOUFFO-TCHEUGUI, J. B. Screening for Diabetes and Prediabetes. **Endocrinology and Metabolism Clinics of North America**, v. 50, n. 3, p. 369–385, set. 2021.

DUNN, J.; RUNGE, R.; SNYDER, M. Wearables and the medical revolution. **Personalized Medicine**, v. 15, n. 5, p. 429–448, set. 2018.

ECHLE, A. *et al.* Deep learning in cancer pathology: a new generation of clinical biomarkers. **British Journal of Cancer**, v. 124, n. 4, p. 686–696, 16 fev. 2021.

FÖLL, S. *et al.* FLIRT: A feature generation toolkit for wearable data. **Computer Methods and Programs in Biomedicine**, v. 212, n. 2021, p. 106461, nov. 2021.

GENG, Z. *et al.* Noninvasive Continuous Glucose Monitoring Using a Multisensor-Based Glucometer and Time Series Analysis. **Scientific Reports**, v. 7, n. 1, p. 12650, 4 out. 2017.

GÉRON, A. **Mãos à Obra Aprendizado de Máquina com Scikit-Learn & TensorFlow: Conceitos, Ferramentas e Técnicas para a Construção de Sistemas Inteligentes**. Rio de Janeiro: Alta Books, 2019.

GHOLAMIANGONABADI, D.; KISELOV, N.; GROLINGER, K. Deep Neural Networks for Human Activity Recognition With Wearable Sensors: Leave-One-Subject-Out Cross-Validation for Model Selection. **IEEE Access**, v. 8, p. 133982–133994, 2020.

GIACAGLIA, L. R. *et al.* Tratamento farmacológico do pré-diabetes. In: **Diretriz Oficial da Sociedade Brasileira de Diabetes**. Porto Alegre: Conectando Pessoas, 2022. p. 1–12.

GOLDBERGER, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet. **Circulation**, v. 101, n. 23, 13 jun. 2000.

GUYTON, A. C.; HALL, J. E. **Tratado de Fisiologia Médica**. 11^a ed. Rio de Janeiro: Elsevier, 2006.

GYŐRFFY, B. Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. **Computational and Structural Biotechnology Journal**, v. 19, p. 4101–4109, 2021.

HEIKES, K. E. *et al.* Diabetes Risk Calculator: A simple tool for detecting undiagnosed diabetes and pre-diabetes. **Diabetes Care**, v. 31, n. 5, p. 1040–1045, 1 maio 2008.

HINA, A.; SAADEH, W. A Noninvasive Glucose Monitoring SoC Based on Single Wavelength Photoplethysmography. **IEEE Transactions on Biomedical Circuits and Systems**, v. 14, n. 3, p. 504–515, jun. 2020.

IBRAHIM, H.; DARWEESH, M. S.; SOLTAN, A. **An Optimized Non-Invasive Blood Glucose and Temperature Body Measurement System**. 2022 11th International

Conference on Modern Circuits and Systems Technologies (MOCAST). **Anais...IEEE**, 8 jun. 2022. Disponível em: <<https://ieeexplore.ieee.org/document/9837536/>>

INTERNATIONAL DIABETES FEDERATION. **IDF Diabetes Atlas**. 10º ed. [s.l.] International Diabetes Federation, 2021.

IQBAL, S. M. A. *et al.* Advances in healthcare wearable devices. **npj Flexible Electronics**, v. 5, n. 1, p. 9, 12 abr. 2021.

JAIN, P.; JOSHI, A. M.; MOHANTY, S. P. iGLU: An Intelligent Device for Accurate Noninvasive Blood Glucose-Level Monitoring in Smart Healthcare. **IEEE Consumer Electronics Magazine**, v. 9, n. 1, p. 35–42, 2020.

JOSHI, A. M. *et al.* iGLU 2.0: A New Wearable for Accurate Non-Invasive Continuous Serum Glucose Measurement in IoMT Framework. **IEEE Transactions on Consumer Electronics**, v. 66, n. 4, p. 327–335, nov. 2020.

KHAN, R. M. M. *et al.* From Pre-Diabetes to Diabetes: Diagnosis, Treatments and Translational Research. **Medicina**, v. 55, n. 9, p. 546, 29 ago. 2019.

KHUNTI, K. *et al.* A comparison of cost per case detected of screening strategies for Type 2 diabetes and impaired glucose regulation: Modelling study. **Diabetes Research and Clinical Practice**, v. 97, n. 3, p. 505–513, 2012.

KONCINA *et al.* Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges. **Cancers**, v. 12, n. 2, p. 319, 30 jan. 2020.

LAWAND, K.; PARIHAR, M.; PATIL, S. N. **Design and development of infrared LED based non invasive blood glucometer**. 2015 Annual IEEE India Conference (INDICON). **Anais...IEEE**, dez. 2015. Disponível em: <<http://ieeexplore.ieee.org/document/7443487/>>

LEHMANN, V. *et al.* Noninvasive Hypoglycemia Detection in People With Diabetes Using Smartwatch Data. **Diabetes Care**, v. 46, n. 5, p. 993–997, 1 maio 2023.

LEVY, M.; PRENTICE, M.; WASS, J. Diabetes insipidus. **BMJ**, p. 1321, 28 fev. 2019.

LIN, T. Non-Invasive Glucose Monitoring: A Review of Challenges and Recent Advances. **Current Trends in Biomedical Engineering & Biosciences**, v. 6, n. 5, p. 1–8, 31 jul. 2017.

NAKAMURA, K.; BALO, A. The Accuracy and Efficacy of the Dexcom G4 Platinum Continuous Glucose Monitoring System. **Journal of Diabetes Science and Technology**, v. 9, n. 5, p. 1021–1026, 1 set. 2015.

OLIVEIRA, G.; MENDONÇA, M. **ExperText: Uma Ferramenta de Combinação de Múltiplos Classificadores Naive Bayes**. Anales de la 4a Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería de Conocimiento. **Anais...2000**.

OLIVEIRA JUNIOR, G. M. DE. **Máquina de Vetores Suporte : estudo e análise de parâmetros para otimização de resultado**. [s.l.] Universidade Federal de Pernambuco, 2010.

ONORATI, F. *et al.* Multicenter clinical assessment of improved wearable multimodal

convulsive seizure detectors. **Epilepsia**, v. 58, n. 11, p. 1870–1879, nov. 2017.

PAULO, M. F. DE; BOCK, E. G. P. **Algoritmo de Detecção da Fibrilação Atrial a Partir dos Intervalos RR**. XXIV Congresso Brasileiro de Engenharia Biomédica. **Anais**...2014.

PEDREGOSA, F. *et al.* Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PICARD, R. W.; BOYER, E. W. Smartwatch biomarkers and the path to clinical use. **Med**, v. 2, n. 7, p. 797–799, 2021.

POSADA-QUINTERO, H. F.; CHON, K. H. Innovations in electrodermal activity data collection and signal processing: A systematic review. **Sensors (Switzerland)**, v. 20, n. 2, 2020.

PRABHA, A. *et al.* **Non-invasive Diabetes Mellitus Detection System using Machine Learning Techniques**. 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence). **Anais**...IEEE, 28 jan. 2021a. Disponível em: <<https://ieeexplore.ieee.org/document/9377138/>>

PRABHA, A. *et al.* Design of intelligent diabetes mellitus detection system using hybrid feature selection based XGBoost classifier. **Computers in Biology and Medicine**, v. 136, n. March, p. 104664, set. 2021b.

PROMPHET, N. *et al.* Cotton thread-based wearable sensor for non-invasive simultaneous diagnosis of diabetes and kidney failure. **Sensors and Actuators B: Chemical**, v. 321, n. June, p. 128549, out. 2020.

QURESHI, F.; KRISHNAN, S. Wearable Hardware Design for the Internet of Medical Things (IoMT). **Sensors**, v. 18, n. 11, p. 3812, 7 nov. 2018.

RAMAZI, R. *et al.* Predicting progression patterns of type 2 diabetes using multi-sensor measurements. **Smart Health**, v. 21, n. June, p. 100206, jul. 2021.

REDA, A. *et al.* Optical glucose biosensor built-in disposable strips and wearable electronic devices. **Biosensors and Bioelectronics**, v. 185, n. April, p. 113237, ago. 2021.

REFARDT, J.; WINZELER, B.; CHRIST-CRAIN, M. Diabetes Insipidus. **Endocrinology and Metabolism Clinics of North America**, v. 49, n. 3, p. 517–531, set. 2020.

RODACKI, M. *et al.* Classificação do diabetes. In: **Diretriz Oficial da Sociedade Brasileira de Diabetes**. Porto Alegre: Conectando Pessoas, 2022. p. 1–28.

RODWELL, V. W. *et al.* **Bioquímica Ilustrada de Harper**. 30^a ed. Porto Alegre: AMGH, 2017.

RUSSEL, S.; NORVIG, P. **Inteligência Artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013.

SANCHEZ, P. *et al.* Real-World Data from the First U.S. Commercial Users of an Implantable Continuous Glucose Sensor. **Diabetes Technology & Therapeutics**, v. 21, n. 12, p. 677–681, 1 dez. 2019.

SCIKIT-LEARN. **User Guide**. Disponível em: <https://scikit-learn.org/stable/user_guide.html>. Acesso em: 11 out. 2022.

SILVA, L.; PATEL, N. **How Much Does A Continuous Glucose Monitor Cost?** Disponível em: <<https://www.forbes.com/health/body/continuous-glucose-monitor-cost/>>. Acesso em: 26 maio. 2023.

SOUZA, C. F. DE *et al.* Pré-diabetes: diagnóstico, avaliação de complicações crônicas e tratamento. **Arquivos Brasileiros de Endocrinologia & Metabologia**, v. 56, n. 5, p. 275–284, jul. 2012.

SUN, H. *et al.* IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. **Diabetes Research and Clinical Practice**, v. 183, p. 109119, jan. 2022.

TAMADA, J. A. *et al.* Noninvasive Glucose Monitoring. **JAMA**, v. 282, n. 19, p. 1839, 17 nov. 1999.

TANG, F. *et al.* Non-Invasive Glucose Measurement by Use of Metabolic Heat Conformation Method. **Sensors**, v. 8, n. 5, p. 3335–3344, 21 maio 2008.

TELLES, S. C. L. *et al.* O Significado Clínico da Actigrafia. **Revista Neurociências**, v. 19, n. 1, p. 153–161, 31 mar. 2001.

TIERNEY, M. J. *et al.* The GlucoWatch® biographer: a frequent, automatic and noninvasive glucose monitor. **Annals of Medicine**, v. 32, n. 9, p. 632–641, 27 jan. 2000.

TIERNEY, M. J. *et al.* Clinical evaluation of the GlucoWatch® biographer: A continual, non-invasive glucose monitor for patients with diabetes. **Biosensors and Bioelectronics**, v. 16, n. 9–12, p. 621–629, 2001.

WENTHOLT, I. M. E. *et al.* Pendra goes Dutch: lessons for the CE mark in Europe. **Diabetologia**, v. 48, n. 6, p. 1055–1058, 4 jun. 2005.

YARMAN, A.; KURBANOGLU, S.; SCHELLER, F. W. Noninvasive biosensors for diagnostic biomarkers. In: **Commercial Biosensors and Their Applications**. [s.l.] Elsevier, 2020. p. 167–181.

YIN, H. *et al.* DiabDeep: Pervasive Diabetes Diagnosis Based on Wearable Medical Sensors and Efficient Neural Networks. **IEEE Transactions on Emerging Topics in Computing**, v. 9, n. 3, p. 1139–1150, 1 jul. 2021.

ZHANG, Y. *et al.* Non-invasive blood glucose detection system based on conservation of energy method. **Physiological Measurement**, v. 38, n. 2, p. 325–342, 1 fev. 2017.