

**UNIVERSIDADE BRASIL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA BIOMÉDICA
CAMPUS SÃO PAULO**

MARCELO HENRIQUE DE VASCONCELOS MOURÃO

**MODELO PREDITIVO DO NÍVEL GLICÊMICO POR
MONITORAMENTO EM TEMPO REAL EM INDIVÍDUOS
PORTADORES DE DIABETES MELLITUS TIPO II**

**PREDICTIVE MODEL OF GLYCEMIC LEVEL BY REAL-TIME
MONITORING IN INDIVIDUALS WITH TYPE II DIABETES MELLITUS**

São Paulo
2022

MARCELO HENRIQUE DE VASCONCELOS MOURÃO

MODELO PREDITIVO DO NÍVEL GLICÊMICO POR
MONITORAMENTO EM TEMPO REAL EM INDIVÍDUOS
PORTADORES DE DIABETES MELLITUS TIPO II

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Biomédica da Universidade Brasil, como parte dos requisitos necessários para obtenção do título de Doutor em Engenharia Biomédica.

Prof. Dr. Marcello Magri Amaral
Orientador

Profa. Dra. Laurita dos Santos
Coorientador(a)

São Paulo - SP

2022

**Ficha catalográfica elaborada pelo Sistema de Bibliotecas da Universidade Brasil,
com os dados fornecidos pelo (a) autor (a).**

M891m MOURÃO, Marcelo Henrique de Vasconcelos.

Modelo preditivo do nível glicêmico por monitoramento em tempo real em indivíduos portadores de diabetes mellitus tipo II / Marcelo Henrique de Vasconcelos Mourão -- São Paulo: Universidade Brasil, 2022.
99 f.: il. color.

Tese de Doutorado defendida no Programa de Pós-graduação do Curso de Engenharia Biomédica da Universidade Brasil.

Orientação: Prof. Dr. Marcello Magri Amaral.

Coorientação: Profa. Dra. Laurita dos Santos

1. Glicemia. 2. Diabetes Mellitus. 3. Inteligência Artificial. 4. Aprendizagem de Máquina. I. Amaral, Marcello Magri. II. Santos, Laurita dos. III. Título.

CDD 610.28



TERMO DE APROVAÇÃO

MARCELO HENRIQUE DE VASCONCELOS MOURÃO

**“MODELO PREDITIVO DO NÍVEL GLICÊMICO POR MONITORAMENTO EM TEMPO
REAL EM INDIVÍDUOS PORTADORES DE DIABETES MELLITUS TIPO II”**

Tese aprovada como requisito parcial para obtenção do título de **Doutor no Programa de Pós-Graduação em Engenharia Biomédica** da Universidade Brasil, pela seguinte banca examinadora:

Prof(a). Dr(a) Marcello Magri Amaral (presidente-orientador)

Prof(a). Dr(a) Livia Assis Garcia (UNIVERSIDADE BRASIL)

Prof(a). Dr(a) Priscila Pereira Fávero (UNIVERSIDADE BRASIL)

Prof(a). Dr(a). Marcus Paulo Rael (INSTITUTO DE PESQUISAS ENERGÉTICAS E NUCLEARES)

Prof(a). Dr(a). Thiago da Silva Cordeiro (JOHN DEERE)

São Paulo, 05 de dezembro de 2022.

Presidente da Banca Prof.(a) Dr.(a) Marcello Magri Amaral

Houve alteração do Título: sim () não (x):



Termo de Autorização

Para Publicação de Dissertações e Teses no Formato Eletrônico na Página WWW do Respectivo Programa da Universidade Brasil e no Banco de Teses da CAPES

Na qualidade de titular(es) dos direitos de autor da publicação, e de acordo com a Portaria CAPES no. 13, de 15 de fevereiro de 2006, autorizo(amos) a Universidade Brasil a disponibilizar através do site <http://www.universidadebrasil.edu.br>, na página do respectivo Programa de Pós-Graduação *Stricto Sensu*, bem como no Banco de Dissertações e Teses da CAPES, através do site <http://bancodeteses.capes.gov.br>, a versão digital do texto integral da Dissertação/Tese abaixo citada, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira.

A utilização do conteúdo deste texto, exclusivamente para fins acadêmicos e científicos, fica condicionada à citação da fonte.

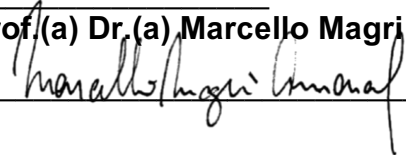
Título do Trabalho: **“MODELO PREDITIVO DO NÍVEL GLICÊMICO POR MONITORAMENTO EM TEMPO REAL EM INDIVÍDUOS PORTADORES DE DIABETES MELLITUS TIPO II”**.

Autor(es):

Discente: **Marcelo Henrique de Vasconcelos Mourão**

Assinatura: 

Orientador(a): **Prof.(a) Dr.(a) Marcello Magri Amaral**

Assinatura: 

Coorientador(a): **Prof.(a) Dr.(a) Laurita dos Santos**

Assinatura: 

Houve alteração do Título: sim () não (x):

Data: 05/12/2022

À minha mãe, por ser fonte de inspiração da origem desta pesquisa, principalmente quando passou a usar o sistema operacional *free style libre*.

AGRADECIMENTOS

A Deus, por ter me guiado até aqui e proporcionado incríveis experiências.

Ao Prof. Dr. Marcello Magrii Amaral, por sua orientação, atenção e cuidado que teve para com esta tese.

À Universidade Brasil (UAB), campus Itaquera-SP, por promover o curso de Doutorado em Engenharia Biomédica.

À Universidade Estadual do Maranhão (UEMA), Campus Santa Inês pelo apoio e incentivo de minha formação e, assim, honrando para com os alunos uma educação de qualidade de acordo a proposta da instituição.

E a todos que direta e indiretamente contribuíram para produção e finalização desta pesquisa.

"Diabético" é quem não consegue ser doce. (Mario Quintana)

RESUMO

Esta pesquisa trata de uma predição dos níveis glicêmicos de pessoas com Diabetes Mellitus II, coletados por meio de um sistema de monitoramento contínuo de glicemia, a partir da arquitetura de redes neurais LSTM. A diabetes, uma das doenças crônicas não transmissíveis, caracteriza-se pela hiperglicemia na corrente sanguínea gerada pela resistência insulínica. O controle dessa doença pode ocorrer por meio da contagem de carboidratos de acordo com o nível glicêmico, que de acordo com a avaliação antropométrica é quantificado pelo médico. Entretanto, essa abordagem nem sempre tem uma boa aceitação pelos diabéticos, que acabam aderindo à medicação para seu controle. Apesar disso, alguns diabéticos acabam por fazer uso de sensores de monitoramento de glicemia contínua, o que favoreceu para verificar se os dados coletados glicêmicos de 15 min em 15 min podiam ser preditos. A glicemia de 20 pacientes foi medida por um período de 14 dias utilizando monitoramento em tempo real. Neste período os hábitos alimentares foram registrados para contagem de carboidratos ingeridos, utilizando o app de contagem de carboidrato criado pela Sociedade Brasileira de Diabetes. Utilizando-se modelo de inteligência artificial (LSTM) foi criado um modelo de predição da glicemia. Com esse modelo verificou-se que os valores preditos acompanharam o movimento glicêmico real, antecipando 5 horas com dados glicêmicos de 12 horas contínuas, ou seja, 20 observações preditas e 48 observações coletadas pelo sensor de glicemia para cada indivíduo. Realizou-se um modelo preditivo geral com os 20 voluntários e dois personalizados. Os dados glicêmicos dos diabéticos coletados tiveram desempenho positivo, pois os valores preditos acompanharam o movimento glicêmico, havendo um pico de glicemia de 170 mg/dL às 9h da manhã e 180 mg/dL às 13h da tarde, convergindo com os dados obtidos da contagem de carboidratos, avaliação física e antropométrica, observados com os picos de glicemia, modos de vida dos voluntários e os carboidratos totais consumidos diariamente. Os dados glicêmicos dos não diabéticos tiveram desempenho positivo, haja vista os dados preditos acompanharem o movimento glicêmico real. Esse modelo, assim, pode predizer várias aplicações direto na reabilitação, contribuindo como um dos instrumentos importantes para a melhora qualidade de vida do paciente.

PALAVRAS-CHAVE: Glicemia. Diabetes Mellitus. Inteligência Artificial. Aprendizagem de Máquina.

ABSTRACT

This research deals with a prediction of the glycemic levels of people with Diabetes Mellitus II, collected through a continuous glycemic monitoring system, based on the architecture of LSTM neural networks. Diabetes, one of the non-communicable chronic diseases, is characterized by hyperglycemia in the bloodstream generated by insulin resistance. The control of this disease can occur through carbohydrate counting according to the glycemic level, which according to the anthropometric evaluation is quantified by the physician. However, this approach is not always well accepted by diabetics, who end up adhering to medication for their control. Despite this, some diabetics end up using continuous blood glucose monitoring sensors, which favored verifying whether the glycemic data collected every 15 minutes could be predicted. The glycemia of 20 patients was measured over a period of 14 days using real-time monitoring. During this period, eating habits were recorded to count ingested carbohydrates, using the carbohydrate counting app created by SBD. Using an artificial intelligence model (LSTM) a blood glucose prediction model was created. With this model, it was verified that the predicted values followed the real glycemic movement, anticipating 5 hours with glycemic data of 12 continuous hours, that is, 20 predicted observations and 48 observations collected by the glycemia sensor for each individual. A general predictive model was performed with 20 volunteers and two personalized ones. The glycemic data of the collected diabetics had a positive performance, as the predicted values followed the glycemic movement, with a glycemic peak of 170 mg/dL at 9 am and 180 mg/dL at 1 pm, converging with the data obtained from the blood count. of carbohydrates, physical and anthropometric evaluation, observed with the peaks of glycemia, lifestyles of the volunteers and the total carbohydrates consumed daily. The glycemic data of non-diabetics had a positive performance, given that the predicted data followed the actual glycemic movement. This model, therefore, can predict several applications directly in rehabilitation, contributing as one of the important instruments for improving the patient's quality of life.

KEYWORDS: Blood glucose. Diabetes Mellitus. Artificial intelligence. Machine Learning.

DIVULGAÇÃO E TRANSFERÊNCIA DE CONHECIMENTO

Esta pesquisa apresenta a inteligência artificial como uma das ferramentas que podem promover uma melhoria no bem-estar das pessoas que possuem diabetes melitus tipo II. Para tanto, contou com 20 voluntários, sendo 10 diabéticos e 10 não diabéticos, os quais responderam formulários e se dispuseram a usar o sensor de monitoramento contínuo, para a obtenção dos dados glicêmicos, medidos de 15 em 15 minutos. Esses dados foram compilados e dispostos na biblioteca TensorFlow e Scikit-learn para a criação e avaliação do modelo LSTM, onde se obteve um resultado positivo quanto à previsão da glicemia dos voluntários em cinco horas.

LISTA DE SIGLAS

ADA	– American Diabetes Association
AGP	– <i>Ambulatory glucose profile</i>
AM	– Aprendizado de Máquina
CGM	– Monitoramento Contínuo de Glicose
DAOP	– Doença Arterial Obstrutiva Periférica
DCNT	– Doenças Crônicas não Transmissíveis
DM	– Diabetes Mellitus
DM1	– Diabetes Tipo 1
DM2	– Diabetes Tipo 2
DMG	– Diabetes Mellitus Gestacional
FDA	– Food and Drug Administration
GJA	– Glicemia de jejum alterada
GMI	– Indicador de gerenciamento de glicose
HA	– Hipertensão Arterial
IA	– Inteligência Artificial
IDF	– Federação Internacional de Diabetes
IMC	– Índice de Massa Corporal
kNN	– k-vizinho mais próximo
LSTM	– Long-Short Term Memory
MAE	– Mean Absolute Error
MFG	– Monitoramento flash da glicose
<i>MLP</i>	– Multi-layer perceptron
MSE	– Mean Squared Error
OMS	– Organização Mundial da Saúde
RMSE	– <i>Root Mean Square Error</i>
RNA	– Redes Neurais Artificiais
RNNs	– Redes neurais recorrentes
SBD	– Sociedade Brasileira de Diabetes
SVM	– Máquinas de vetor de suporte
TCLE	– Termo de compromisso livre esclarecido
TDG	– Tolerância diminuída à glicose
VG	– Verificação do intervalo glicêmico

LISTA DE FIGURAS

Figura 01 –	Sistema de monitoramento de glicose intersticial	25
Figura 02 –	Exemplo de gráfico de padrões diários (AGP)	27
Figura 03 –	Exemplo de informações retrospectivas	28
Figura 04 –	Representação gráfica de uma rede perceptron	32
Figura 05 –	Representação gráfica de uma rede perceptron multi-camadas	32
Figura 06 –	Representação gráfica de uma rede neural recorrente	33
Figura 07 –	Arquitetura de uma rede LSTM	34
Figura 08 –	Método <i>Holdout</i>	35
Figura 09 –	Método <i>k-Fold</i>	36
Figura 10 –	O tamanho do sensor do freestyle libre x moeda de 1 centavo americano	38
Figura 11 –	Glicose média ao longo das observações dos dados dos voluntários diabéticos	45
Figura 12 –	Glicose média ao longo das observações dos dados dos voluntários não diabéticos	45
Figura 13 –	Sexo dos voluntários	49
Figura 14 –	Faixa etária dos voluntários	49
Figura 15 –	Faixa salarial dos voluntários	50
Figura 16 –	Formação educacional salarial dos voluntários DB	50
Figura 17 –	Formação educacional salarial dos voluntários NDB	51
Figura 18 –	Atividades dos voluntários	51
Figura 19 –	Consumo de bebida alcoólica e tabagismo dos voluntários	52
Figura 20 –	Patologias e Intervenções Cirúrgicas dos voluntários	52
Figura 21 –	Peso dos voluntários	54
Figura 22 –	Estatura dos voluntários	54
Figura 23 –	Classificação do IMC de acordo com a OMS dos voluntários	55
Figura 24 –	Contagem de Carboidrato dos voluntários diabéticos	55
Figura 25 –	Contagem de Carboidrato dos voluntários não diabéticos	56
Figura 26 –	Resultado do modelo LSTM treinado com todos os dados	58
Figura 27 –	Resultado do modelo LSTM treinado com os dados dos voluntários diabéticos	60
Figura 28 –	Resultado do modelo LSTM treinado com os dados dos voluntários não diabéticos	60

LISTA DE TABELAS

Tabela 01 –	Interpretação das setas de tendência, advinda do SMCG	29
Tabela 02 –	Classificação da pressão arterial em adultos	41
Tabela 03 –	Classificação do IMC de acordo com a OMS	42
Tabela 04 –	Avaliação Física dos sinais vitais dos voluntários diabéticos	53
Tabela 05 –	Avaliação Física dos sinais vitais dos voluntários não diabéticos	53
Tabela 06 –	Descrição do modelo LSTM com todos os voluntários	56
Tabela 07 –	Resultado da validação cruzada dos dados, contendo os voluntários diabéticos e não diabéticos	57
Tabela 08 –	Métricas de Regressão	57
Tabela 09 –	Descrição dos modelos LSTM treinados individualmente com os dados de um voluntário diabético	59
Tabela 10 –	Descrição dos modelos LSTM treinados individualmente com os dados de um voluntário não diabético	59
Tabela 11 –	Modelo dos voluntários diabéticos	59
Tabela 12 –	Modelo dos voluntários não diabéticos	59

LISTA DE SIGLAS

ADA	American Diabetes Association
AGP	<i>Ambulatory glucose profile</i>
AM	Aprendizado de Máquina
CGM	Monitoramento Contínuo de Glicose
DAOP	Doença Arterial Obstrutiva Periférica
DCNT	Doenças Crônicas não Transmissíveis
DM	Diabetes Mellitus
DM1	Diabetes Tipo 1
DM2	Diabetes Tipo 2
DMG	Diabetes Mellitus Gestacional
FDA	Food and Drug Administration
GJA	Glicemia de jejum alterada
GMI	Indicador de gerenciamento de glicose
HA	Hipertensão Arterial
IA	Inteligência Artificial
IDF	Federação Internacional de Diabetes
IMC	Índice de Massa Corporal
kNN	k-vizinho mais próximo
LSTM	Long-Short Term Memory
MAE	Mean Absolute Error
MFG	Monitoramento flash da glicose
<i>MLP</i>	Multi-layer perceptron
MSE	Mean Squared Error
OMS	Organização Mundial da Saúde
RMSE	<i>Root Mean Square Error</i>
RNA	Redes Neurais Artificiais
RNNs	Redes neurais recorrentes
SBD	Sociedade Brasileira de Diabetes
SVM	Máquinas de vetor de suporte
TCLE	Termo de compromisso livre esclarecido
TDG	Tolerância diminuída à glicose
VG	Verificação do intervalo glicêmico

SUMÁRIO

1 INTRODUÇÃO	15
2 OBJETIVOS	20
2.1 OBJETIVO GERAL	20
2.2 OBJETIVOS ESPECÍFICOS	20
3 REFERENCIAL TEÓRICO	21
3.1 SISTEMA DE MONITORAMENTO CONTÍNUO DE GLICOSE	24
3.1.1 Setas de tendência	29
3.2 APRENDIZADO DE MÁQUINA.....	29
3.3 ALGORITMOS DE APRENDIZADO SUPERVISIONADO	31
3.4 REDES NEURAIS ARTIFICIAIS (RNA)	31
3.4.1 Perceptron	31
3.4.2 Perceptron Multi-camadas	32
3.4.3 Redes Neurais Recorrentes	33
3.5 VALIDAÇÃO CRUZADA	35
3.5.1 Holdout	35
3.5.2 k-Fold	36
4 MATERIAIS E MÉTODOS	37
4.1 CONSIDERAÇÕES ÉTICAS	37
4.2 TIPO DE ESTUDO: DESENHO	38
4.3 LOCAL DE COLETA DE DADOS.....	38
4.4 PARTICIPANTES.....	39
4.4.1 Caracterização sociodemográfica e de hábitos de vida	40
4.5 INSTRUMENTOS UTILIZADOS PARA COLETA DE DADOS	40
4.5.1 Avaliação Física	40
4.5.2 Avaliação Antropométrica	42
4.5.3 Contagem de Carboidrato	43
4.6 METODOLOGIA DE ANÁLISE DOS DADOS	43
4.6.1 Processamento dos dados	44
4.7 CRIAÇÃO E TREINAMENTO DO MODELO LSTM	46
4.8 MODELO PERSONALIZADO	48
5 RESULTADOS	49
5.1 PREDIÇÃO DOS DADOS SISTÊMICOS	56
6 DISCUSSÕES	61
7 CONCLUSÃO	64
REFERÊNCIAS	66
APÊNDICES	71

1 INTRODUÇÃO

O Diabetes Mellitus (DM), uma das quatro doenças crônicas não transmissíveis, é considerada pela OMS (Organização Mundial de Saúde) e pelo Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas não Transmissíveis (DCNT), 2011-2022, como prioritária para intervenção. O DM é visto assim como um problema de saúde pública mundial. (MALTA, 2019)

A Federação Internacional de Diabetes (IDF) estimou que 8,7% de brasileiros entre a faixa etária de 21 a 80 anos conviviam com diabetes mellitus, sendo que, se as projeções se mantiverem, acarretará 9,1%, um aumento na mesma faixa etária (IDF, 2018), o que o torna ameaçador à vida das pessoas por se tratar de uma doença crônica e que acarreta várias complicações, doenças cardiovasculares, nefropatia, retinopatia, neuropatia, todas sendo problemas consequentes do diabetes Mellitus.

O agravamento da doença está correlacionado especialmente ao envelhecimento da população, falta de cuidados preventivos, má alimentação e baixo nível de exercício físico, sedentarismo, obesidade, consumo de álcool e tabaco, além de fatores econômicos (ZIMMET et al, 2007 apud LYRA, CAVALCANTI, SANTOS, 2019). O diabético necessita de cuidados especiais que requer decisões por especialistas sobre o uso de medicação, se necessário o uso de insulina, e monitorização sanguínea, dietas, atividade física entre outras. (SBD, 2019).

O Diabetes Mellitus (DM) é uma doença caracterizada pela hiperglicemia. O acompanhamento dos níveis glicêmicos é fundamental para o tratamento do DM. Em casos avançados da doença, o paciente deve fazer uso da insulinoterapia para o controle da doença. O não tratamento e acompanhamento dos níveis glicêmicos em portadores de DM podem levar a complicações, disfunções e até ao óbito. (BERTONHI; DIAS, 2019)

O DM pode ser causado por diversos mecanismos que levam à desregulação da produção da insulina – hormônio secretado pelo pâncreas, órgão que produz insulina e transporta a glicose para dentro das células, sendo convertida em energia e utilizada imediatamente, ou armazenada em forma de glicogênio ou gordura. Quando o organismo não produz a quantidade de insulina para o devido transporte,

as células não respondem normalmente a insulina, criando a resistência à insulina e a elevação dos níveis glicêmicos no sangue, complicando a vida do diabético. Os tipos mais comuns de DM são: Diabetes Tipo 1 (DM1), ou Diabetes Juvenil, e Diabetes Tipo 2 (DM2). O DM1 corresponde a 10 %, enquanto o DM2 corresponde a 90 % dos casos. Existe ainda o diabetes gestacional que, diagnosticado durante a gestação, pode levar a perda e má formação fetal, além de apresentar risco à gestante. (SBD, 2019).

Observa-se que a idade e o nível socioeconômico podem influenciar no tratamento contínuo do paciente. É necessário conhecer as complicações individuais para retardar complicações severas, o uso de medidores de glicemia e uma consulta ao especialista, tornando-se indispensáveis para o favorecimento do tratamento. Assim, o uso do sistema de monitoramento contínuo de glicose proporciona informações relevantes ao momento da consulta. (SOUZA, 2017).

O monitoramento da glicemia é uma grande aliada dos portadores de DM e profissionais da saúde. Há mais de quatro décadas no mercado, as fitas para medição da glicose no sangue capilar permitem a aferição do nível glicêmico imediato. Com o avanço das tecnologias, os dispositivos para o prognóstico da glicemia (os glicosímetros) passaram por um processo de aperfeiçoamento e barateamento e, assim, tornaram-se imprescindíveis para o controle do DM. (ROSSANEIS et al., 2020).

Este tipo de tecnologia é minimamente invasivo e necessita de uma ação direta do usuário para medição do nível glicêmico e retirada de uma pequena quantidade de sangue para realizar a medição, geralmente da ponta dos dedos ou lóbulo auricular. Assim, os níveis glicêmicos são adquiridos somente em momentos específicos do dia. Isso pode limitar a avaliação, dependendo da frequência e disponibilidade do teste pelo paciente.

Em 2015, a FDA (*Food and Drug Administration*) dos Estados Unidos aprovou o sistema de monitoramento contínuo de glicose (CGM) para uso pessoal em pacientes com diabetes: o sensor descartável aplicado na parte de trás do braço com vida útil de 14 dias, aferindo a glicemia de 15 em 15 minutos automaticamente. Com isso, o paciente pode usar o leitor e aferir quando necessário.

Com os dados de níveis glicêmicos pessoas com diabetes e sem diabetes, a Inteligência Artificial tornou possível o campo da medicina de precisão propenso a

verificar a predição da glicemia contribuindo para a saúde pública e, conseqüentemente, para a qualidade de vida das pessoas com diabetes.

Essa possibilidade de relacionar a Inteligência Artificial (IA) com a medicina levou Berry et al (2020) a desenvolver um modelo de aprendizado de máquina que previu as respostas de triglicerídeos e glicêmicas à ingestão de alimentos, utilizando modelos de Regressão de Floresta Aleatória ajustados com os seguintes recursos informativos (composição da refeição, dieta habitual, contexto da refeição, antropometria, genética, microbioma, parâmetros clínicos e bioquímicos) para prever triglicerídeos, glicose e peptídeo C em um conjunto de dados do Reino Unido.

Machine Learning, Aprendizado de Máquina, é uma das áreas da Inteligência Artificial que tomou forma nas últimas décadas, sendo utilizado em vários sistemas de medicina, especialmente na tomada de decisão clínica (MATHUR et al, 2022). O objetivo do aprendizado de máquina na presente pesquisa visa entender a estrutura dos dados e encaixá-los em modelos que as pessoas possam entender e usar.

Outro exemplo de sucesso com os principais métodos de Aprendizado de Máquina correlacionados com a área da saúde encontra-se na pesquisa de Tanveer et al (2020). Os resultados mostram que os modelos baseados em Máquinas de Vetor de Suporte (SVM) têm sido amplamente utilizados para a doença de Alzheimer mostrando sua robustez. As técnicas de Redes Neurais Artificiais (RNA) profundas fornecem resultados promissores ao modelar dados altamente complexos com alta precisão. Mas o grande uso do SVM também decorre do fato de ser mais fácil de interpretar em comparação com as redes neurais profundas, que atuam como modelos de caixa preta.

Apesar de haver outras abordagens como o sistema Arima e o Profeta, escolheu-se o LSTM por ter os dados coletados uma característica serial, ao longo de um determinado intervalo de tempo, adequando-se ao propósito da pesquisa: na resolução de problemas com dados temporais. Além disso, Wang e Li (2020) explicam que o aprendizado profundo, especialmente a memória de longo prazo (LSTM), é mais adequado em estudos de glicemia, pois:

Os estudos de glicemia existentes são comprovadamente eficazes apenas em 30 minutos, mas a precisão se deteriora drasticamente quando o horizonte de previsão aumenta para 45 minutos e 60 minutos. O aprendizado profundo, especialmente para memória de longo prazo (LSTM) e suas variantes, foram recentemente aplicados em várias áreas para alcançar resultados de ponta em tarefas com dados complexos de séries temporais. (WANG; LI, 2020, p. 3256)

Diante disso, surge a seguinte questão: Qual a melhor forma de se obter uma predição de nível glicêmico dos sujeitos que possuem diabetes mellitus tipo II? O modelo LSTM surge como uma das ferramentas que melhor pode prever o nível glicêmico, devido ao grau de precisão com os dados temporais, de acordo com os estudos de Wang e Li, pois a precisão se deteriora drasticamente quando o horizonte da previsão aumenta para 45 min ou 60 min com outras ferramentas. Mas observou-se que o aprendizado profundo, como a memória de longo prazo (LSTM), torna-se eficiente para os tipos de dados glicêmicos com uma variabilidade de tempo maior.

Por isso, ao observar que as medições do nível glicêmico podem ser visualizadas pelo visor do leitor¹ do aparelho com sistema de monitoramento contínuo de glicose que também apresenta setas de tendências sobre o percentual glicêmico não o torna suficiente para uma possível predição, mas somente para medição, o que possibilita apenas uma indicação quando o nível glicêmico se encontra em tendência a subir ou descer em diversos níveis. Entenda-se que essa observação se torna possível para aqueles que atuam na área da saúde que trabalham com o aparelho de forma interpretativa, mas para usuários comuns e leigos sobre o assunto isso torna-se quase impossível.

Esta pesquisa surge assim com o objetivo de desenvolver um método de predição, utilizando técnicas de Inteligência Artificial a partir de séries temporais dos níveis glicêmicos obtidos por tecnologia CGM, aplicados em pacientes diabéticos tipo II.

Assim, a escolha do tema exposto deve-se a uma constatação de que, no paciente com diabetes tipo II que fazia uso do aparelho com sistema de monitoramento contínuo de glicose, havia possibilidades para uma aferição futura da glicemia. Nesse leitor CGM, há a presença de uma seta de tendência que indica o aumento ou a baixa de glicemia. Entretanto, a ausência de aviso sonoro das setas de tendência instigou uma análise dos dados temporais por meio da Inteligência Artificial com o intuito de prever o nível glicêmico dos diabéticos tipo II.

¹ Entende-se por *leitor* o registro no visor que o aparelho realiza ao medir com o sensor o resultado da glicemia no momento solicitado. Não se deve confundir leitor com o *display* do aparelho nem com o sensor que faz a medição, ou seja, compreende-se leitor aqui com o resultado obtido no momento solicitado da glicemia.

Essa pesquisa contribuirá não somente para a academia, mas também para uma melhor qualidade de vida para pessoas com diabetes melitus 2, a fim de neutralizar desconfortos que a hipo e a hiperglicemia provoca no indivíduo ao anteceder cinco horas desses eventos, tempo necessário para uma possível intervenção seja alimentar ou medicamentosa.

A presente pesquisa torna-se original quanto ao tema, pois os estudos que envolvem a inteligência artificial e o diabetes melitus II não buscam prever a glicemia em 5 horas e a finalidade se diferencia da pesquisa aqui em questão. O artigo científico *Inteligência Artificial: o futuro do tratamento do diabetes* de Samer Ellahham não mostra como os dados são tratados nem como isso pode prever os dados glicêmicos, apenas cita alguns dispositivos de inteligência artificial no tratamento da diabete mellitus. Já o artigo *Previsão de glicose no sangue usando variantes LSTM no contexto do sistema de pâncreas artificial de código aberto* de Wang e Li, faz uma análise comparativa das mensurações com LSTM de glicemia partindo dos dados obtidos pela insulina produzida pelo pâncreas.

Este trabalho apresenta, no primeiro momento, a revisão de literatura que explica o contexto da medição do Diabetes Mellitus Tipo II; o uso do sensor com sistema de monitoramento contínuo da glicose; e o uso de inteligência artificial para predição de dados glicêmicos. Em seguida, descreve os dados coletados do uso do CGM em 10 voluntários com diabetes mellitus tipo II e 10 voluntários não diabéticos, da avaliação física e contagem dos carboidratos, do mapeamento da situação do participante e a caracterização dos participantes; e explicação da metodologia com Aprendizagem de Máquina e a linguagem Python para obtenção dos processos de dados nos níveis glicêmicos aferidos e compilados, criando um modelo sistêmico. E, nos resultados, a da criação de uma programação com o uso de inteligência artificial, uma predição da glicemia de diabéticos tipo II por meio de análises em série dos dados coletados de vários voluntários pelo sistema de CGM.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Este projeto teve o objetivo de desenvolver um método de predição do nível glicêmico utilizando técnicas de Inteligência Artificial (IA) a partir de séries temporais obtidos por sistema de monitoramento contínuo de glicose.

2.2 OBJETIVOS ESPECÍFICOS

- Obter as séries temporais dos níveis glicêmicos de indivíduos portadores de DM2;
- Avaliar diferentes abordagens para análise de séries temporais por IA;
- Desenvolver método de predição do nível glicêmico por IA.

3 REFERENCIAL TEÓRICO

O Diabetes Mellitus se caracteriza por um grupo de doenças metabólicas que geram a hiperglicemia, resultante de defeitos na secreção e/ou ação insulínica e, geralmente, acompanhada por dislipidemia, hipertensão arterial (HA) e disfunção endotelial. (SBD, 2019; ELLAHHAM, 2020).

Segundo estudos da Sociedade Brasileira de Diabetes (2019), no diabetes mellitus, há variações no desenvolvimento e no tipo devido a uma série de mecanismos patogênicos (desde a destruição autoimune das células betapancreáticas até anormalidades na ação ou na resistência periférica à ação da insulina). Tem-se como principais sintomas decorrentes de hiperglicemia acentuada: perda de peso (algumas vezes com polifagia), poliúria e polidipsia. Além disso, há uma predisposição a determinadas infecções e sua recorrência.

Nos últimos anos, o diabetes mellitus tem aumentado significativamente, de forma epidêmica, em diferentes regiões do mundo. O DM ainda gera gastos expressivos em saúde, além de provocar uma redução considerada da capacidade de trabalho e da expectativa de vida, o que leva ao aumento do uso dos serviços de saúde. A maioria dos países gasta em média entre 5% e 20% de gastos orçamentários destinados à saúde pública com DM. (IDF, 2017; ROSA et al, 2018)

O Brasil ocupa o 4º lugar dentre os países com maior incidência de DM e afeta um contingente de 3% da população mundial. Em 2017, estimava-se um total de 415 milhões de pessoas diabéticas no mundo, podendo chegar em 2040 a 645 milhões (IDF, 2017). Diante desse contexto, o DM tornar-se motivo de preocupação:

O diabetes mellitus (DM) é uma doença crônica cada vez mais presente, de crescente importância na saúde pública. Segundo dados da International Diabetes Federation (IDF), essa condição afeta aproximadamente 425 milhões de pessoas entre 20 e 79 anos de idade em 2017, com projeção alarmante para 629 milhões de afetados em 2045. (LYRA; CAVALCANTI; SANTOS, 2019, p.18)

Estimativas apontam que, em 2025, a incidência de cerca de 11 milhões de diabéticos. Na maioria dos países, mais de 50% desconhecem a presença do diabetes e dos 20% a 30% que conhecem, poucos fazem adesão ao tratamento. (IDF, 2017)

Por isso, torna-se necessário compreender o mecanismo do diabetes mellitus, assim como auxiliar as pessoas no processo de adesão ao tratamento e ao

autocuidado. Entende-se como autocuidado: o consumo alimentar, a atividade física regular, o uso da terapia medicamentosa e o acompanhamento de saúde regular. Essas ações são consideradas estratégias imprescindíveis para o controle do DM e suas complicações.

Os vários tipos de diabetes (DM1, DM2, diabetes gestacional e outros tipos) apresentam uma genética molecular específica que envolve o metabolismo da glicose, desenvolvimento, controle e função das células pancreáticas. Entretanto, as investigações neste campo continuam a ser pesquisadas, haja vista o mecanismo do desenvolvimento e suas complicações não serem plenamente compreendidos. (SBD, 2019; ELLAHHAM, 2020)

A classificação respalda-se no conceito de estágios clínicos da doença, que incluem a normoglicemia, a tolerância diminuída à glicose (TDG), a glicemia de jejum alterada (GJA) e até o quadro de diabetes *per se*. A Organização Mundial de Saúde (OMS), American Diabetes Association (ADA) e a Sociedade Brasileira de Diabetes (SBD) propõem a seguinte classificação, incluindo as quatro classes clínicas: DM tipo 1, DM tipo 2 (DM2), outros tipos específicos DM e diabetes mellitus gestacional (DMG), sendo os mais recorrentes DM1 e DM2. Assim, tem-se as seguintes características de cada tipo de Diabetes:

- **Tipo 1:** destruição da célula beta, o que ocasiona geralmente deficiência absoluta de insulina, de natureza autoimune ou idiopática.
- **Tipo 2:** ocorre entre a predominância de resistência insulínica com relativa deficiência de insulina e a um defeito essencialmente secretório, com ou sem resistência insulínica.
- **Diabetes gestacional:** sem diagnóstico de diabetes prévio à gestação, é diagnosticado entre o segundo e o terceiro trimestre.
- **Outros tipos específicos:** Doenças do pâncreas exócrino; Endocrinopatias; induzidos por fármacos e agentes químicos; Infecções; Diabetes neonatal; Síndromes diabéticas monogênicas (MODY).

O diagnóstico do DM2, na maioria dos casos, dá-se a partir dos 35 a 40 anos, apesar de ocorrer precocemente durante a adolescência, e é responsável por 90% de todos os casos de diabetes. Observa-se que a maioria dos diabéticos apresenta sobrepeso ou obesidade, o que por si só provoca resistência insulínica; os diabéticos não obesos, ao apresentar deposição elevada de gordura abdominal,

levam a maior propensão para resistência à insulina. (LYRA; CARVALHO; COELHO, 2019)

O estado hiperosmolar não cetótico é uma das complicações de maior incidência nos diabéticos; mas, em casos de infecções graves, há grande propensão para a ocorrência de cetoacidose diabética.

As pessoas com Diabetes Mellitus tipo 2 não necessitam inicialmente de insulina para controlar a glicemia, pois há diferentes causas da doença e não ocorre destruição autoimune completa das células betapancreáticas. (SBD, 2019; ELLAHHAM, 2020)

Esse tipo de diabetes pode ocorrer sem ser diagnosticado por vários anos, pois sua evolução é lenta e gradual, com sintomas leves e até mesmo imperceptíveis. Por isso, os indivíduos que desenvolvem o DM2 “apresentam risco elevado para o desenvolvimento de complicações micro e macrovasculares” (SBD, 2019, p. 20). Observa-se, assim, que o diabete mellitus tipo 2 encontra-se, geralmente, em pacientes com sobrepeso ou obesidade, com histórico da doença na família, e em indivíduos com hipertensão ou dislipidemia. Segundo a Sociedade Brasileira de Diabetes (2019, p. 20), tem-se os seguintes fatores de risco para o desenvolvimento desse tipo:

- Pré-diabetes;
- Raça/etnia de alto risco para DM (negros, hispânicos ou índios Pima);
- História de doença cardiovascular;
- Colesterol HDL < 35 mg/dL e/ou triglicérides > 250 mg/dL;
- Síndrome de ovários policísticos;
- Acantose nigricans;
- Idade \geq 45 anos;
- História familiar de DM (pais, filhos e irmãos);
- Excesso de peso (índice de massa corporal [IMC] \geq 25 kg/m²);
- Sedentarismo;
- HA;
- DMG prévio;
- Macrossomia ou história de abortos de repetição ou mortalidade perinatal;
- Uso de fármacos hiperglicemiantes (por exemplo, corticosteroides, tiazídicos, betabloqueadores).

Em resumo, a genética do Diabetes Mellitus tipo 2 não é claramente definida, por isso complexa. Apesar de haver fatores de risco que favoreçam o desenvolvimento desse tipo, mas não são causas determinantes na presença dele.

3.1 SISTEMA DE MONITORAMENTO CONTÍNUO DE GLICOSE

O monitoramento contínuo da glicose (MCG) demonstra em tempo decisivo a taxa glicêmica, obtendo uma maior quantidade de dados, advertindo para hipoglicemia ou hiperglicemia. Não há necessidade de calibração para oportunizar resultados confiáveis, pois já vem assim disposto de fábrica.

Em 2015, no Congresso da Associação Americana de Diabetes, uma nova metodologia foi lançada para substituir a desconfortável picadinha (ALBUQUERQUE, 2022): o MCG, um sensor colocado na camada subcutânea. Com o monitoramento contínuo da glicose (MCG), uma simples aproximação do aparelho sobre o sensor revela o valor atual imediato, podendo ser aferido, por diversas vezes, diariamente, onde o leitor, além da informação da glicemia, gera dados para uma melhor leitura e uma seta de tendência, que indica a conduta da glicose. Apesar da praticidade e conforto proporcionados pelo MCG, o equipamento não possui alarmes que possam avisar o paciente para equilibrar seus dados glicêmicos.

O diabetes é uma doença silenciosa e os seus portadores estão sujeitos a complicações, disfunções e escassez de vários órgãos, podendo chegar a óbito. Entendemos que os avanços tecnológicos aplicados à Saúde podem ser considerados elementos facilitadores para a prevenção, diagnóstico, acompanhamento e tratamentos adequados, reduzindo a ocorrência de procedimentos invasivos e agilizando exames e resultados.

Neste cenário, uma análise preditiva de possíveis oscilações da taxa glicêmica de portadores de Diabetes Mellitus tipo 2 através do processamento de séries temporais obtidos por monitoramento contínuo de glicose se faz necessário. Assim, os dados glicêmicos foram performados por um modelo supervisionado desenvolvido com linguagem de programação Python na plataforma Jupyter.

A monitorização glicêmica com frequência e regularidade torna-se uma ferramenta essencial principalmente para aqueles que se encontram em terapia

insulínica. O CGM, segundo Fokkert et al. (2017), promove à curto e longo prazo melhoras na condição de saúde.

Apesar da medição de glicose no sangue ser o mais difundido, o automonitoramento vem se tornando habitual e comum no estilo de vida de vários diabéticos. Essa possibilidade originou-se devido, principalmente, às dores sentidas na realização dos testes como principal motivo de queixa. Segundo Hoss e Budiman (2017), os sistemas de monitoramento contínuo de glicose que medem os níveis de glicemia por meio do líquido intersticial surgiram a uns 08 anos.

Figura 01 – Sistema de monitoramento de glicose intersticial. **(A)** Sensor e aparelho de leitura da glicose; **(B)** aplicação do sensor com dispositivo específico do sistema; **(C)** leitura da glicose – passagem do leitor sobre o sensor; **(D)** glicemia medida do paciente.



Fonte: (STEFENON, 2019)

Esse sistema funciona por meio de um sensor colocado na lateral do braço, a partir dos quatro anos de idade, incluindo mulheres grávidas, substituindo a automonitorização da glicemia capilar na autogestão da doença, e utilizado durante um período máximo de 14 dias, sem necessitar de calibração, a partir da tecnologia *Wired enzyme*. (HAAK et al, 2017). O software *FreeStyle Librelink* é gratuito, online e fornece os dados aferidos em txt e pdf.

A medição pode ser feita a cada minuto, em tempo real, bastando somente passar o leitor sobre o sensor. O sensor consiste em uma ponta pequena, flexível, que é inserida logo abaixo da pele quando aplicada na parte lateral do braço do usuário. E os dados são armazenados a cada 15 minutos após aplicado. No visor também se visualiza a tendência de glicose, que indica a mudança do índice glicêmicos que o indivíduo está sujeito, ou seja, “uma seta de tendência que indica a direção e intensidade da variação prevista da glicemia” (CARDOSO et al, 2018, p. 141).

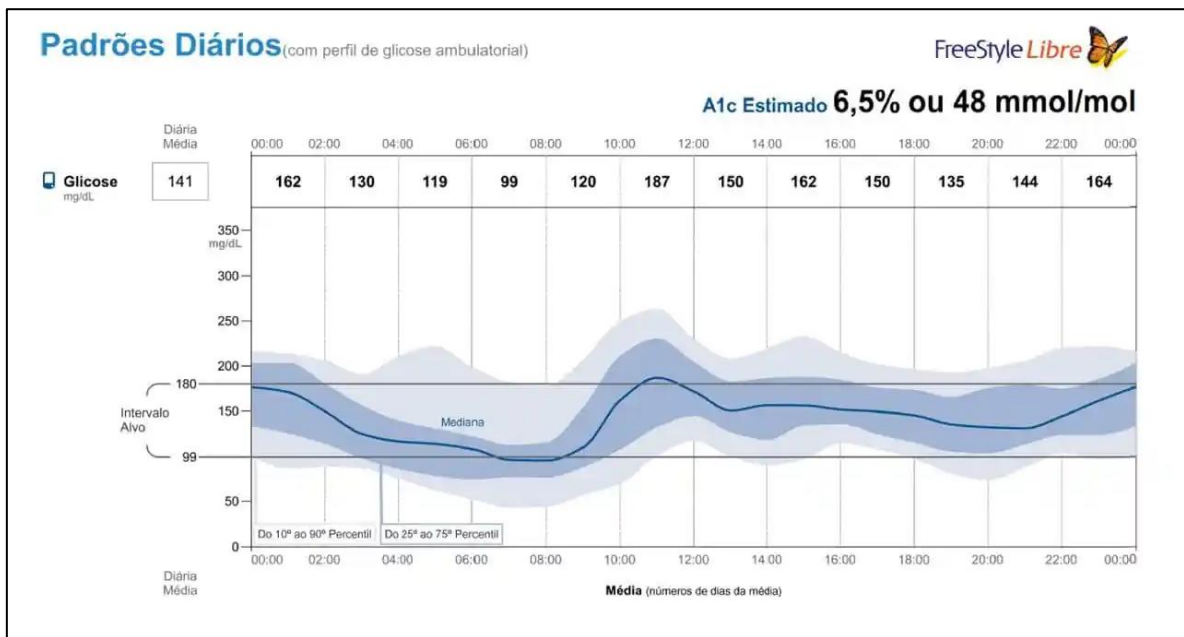
Não existe relatos de interferência medicamentosa entre a aferição do sensor, isto é, não fornece leituras duvidosas. A vantagem do uso desse equipamento é que seus dados podem ser enviados para o médico ou profissionais de saúde que acompanham o paciente, para a melhor decisão sobre medicamentos e suas dosagens.

Para melhor visualização e acompanhamento do nível glicêmico, um gráfico é gerado para um retroativo de 8 horas, o que faz com que o usuário não ultrapasse esse período para fazer a leitura da glicose e, assim, não haja possibilidade de perda de dados. Caso não use o sensor durante o intervalo mínimo de 8 horas, pode parecer uma perda de dados (HOSS et al, 2014; HEINEMANN, FRECKMANN, 2015). Vale ressaltar, no dispositivo para leitura, fica armazenado todas as coletas desde o primeiro dia de uso, quando manuseado de forma adequada.

O programa tecnológico do FreeStyle Libre® (2014) gera o relatório do AGP² que se encontra em conformidade com o consenso internacional e recomendado como padrão para a análise dos dados da glicose no processo de decisão clínica em pacientes com DMT1 e DMT2. (CARDOSO et al, 2018, p. 141).

² Ambulatory Glucose Profile (Perfil Ambulatorial de Glicose)

Figura 02 – Exemplo de gráfico de padrões diários (AGP)

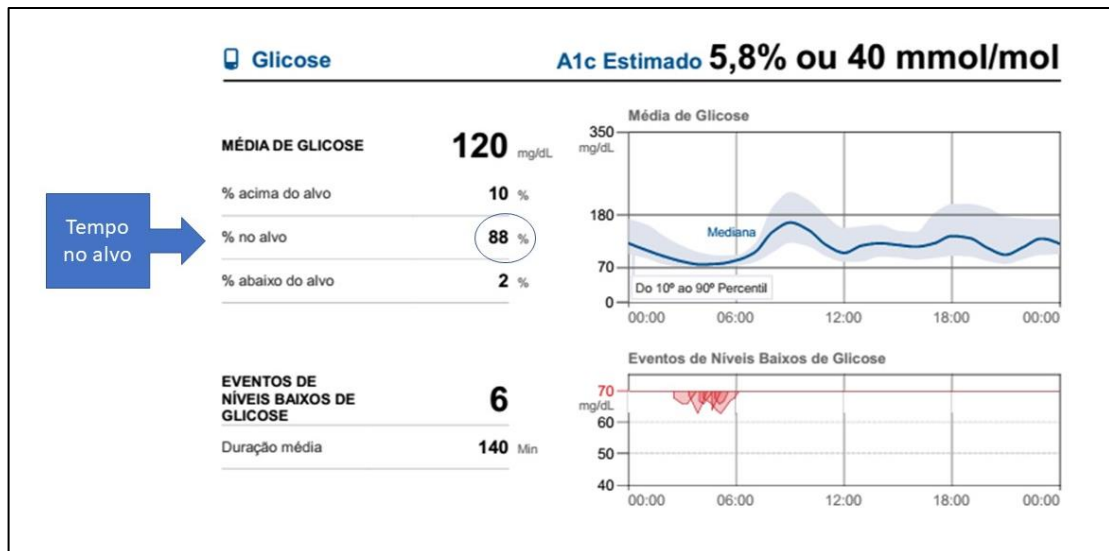


Fonte: VIEIRA, 2019

Assim, o leitor comunica-se com o computador por um software gratuito no qual se forma o gráfico do perfil ambulatorial de glicose (*ambulatory glucose profile*, AGP). O AGP é o padrão-ouro de gráfico para a análise da média glicêmica, da VG (verificada pelo intervalo interquartil de 25^o a 75^o), dos riscos de hipoglicemia por horário do dia, dos níveis de glicose abaixo da meta inserida (quantidade de vezes e duração) e das glicemias antes e depois das refeições. Os dados fornecidos pelo AGP e pelo leitor, juntos, oferecem uma gama potencialmente esmagadora de métricas de glicose no sangue. (SBD, 2019)

Atualmente, os dados do leitor ou do aplicativo podem ser baixados (download) para o celular (para os escaneamentos – LibreLink®) em uma plataforma compartilhada entre pacientes e profissionais de saúde: LibreView®, onde há possibilidade de se analisar o coeficiente de variabilidade e desvio-padrão e comparação com períodos anteriores (14, 30, 60 dias ou personalizados), inexistentes no software inicial. (SBD, 2019)

Figura 03 – Exemplo de informações retrospectivas



Fonte: VIEIRA, 2019

As setas de tendência fornecem informações prospectivas. As informações retrospectivas incluem “indicador de gerenciamento de glicose” (GMI ou antiga HbA1c estimada) de 7, 14 e 30 dias anteriores, VG (variação de glicemia) em formato de gráfico com registros de média de desvios-padrão, média de glicose, proporção de valores acima, abaixo ou na meta de 70 a 100 mg/dL (tempo no alvo), número e duração dos eventos hipoglicêmicos, número de escaneamentos e assim por diante. (CHAWLA, 2019; SBD, 2019).

Tem-se, como característica do sistema de monitoramento contínuo de glicose, uma alternativa a outros métodos de monitoramento da glicemia, o atraso das alterações de glicose no líquido intersticial em relação as alterações da glicose no sangue, o que se considera perfeitamente normal por se tratar de meios diferentes. (FREESTYLE, 2014; BAILEY, 2015).

Há situações que a glicose capilar e intersticial apresenta menor correlação, durante as primeiras 24 horas de uso do CGM; em hipoglicemias e hiperglicemias; e quando mudanças rápidas de glicose ocorrem (período pós-prandial, exercícios). Orienta-se nesses casos a análise da glicemia capilar pela AMGC, dando preferência para este último dado. (SBD, 2019; HEINEMANN, FRECKMANN, 2015)



Essa tecnologia mostra-se eficaz e segura, principalmente, em indivíduos adultos com DM1 e DM2 nos quais as pesquisas mais se voltam como objetos de estudo. Entretanto, não há nenhum estudo que comprove a redução das

complicações da doença, no caso a mortalidade, apesar de reduzir a duração de hipo e hiperglicemia. (SBD, 2019; CHICO et al., 2019)

3.1.1 Setas de tendência

O dispositivo com CGM apresenta uma vantagem importante: as setas de tendência direcional (fornecidas ao lado da leitura atual da glicose). As setas dão informações sobre a direção e a taxa de mudança (rate of change) dos níveis de glicose intersticial. E são geradas 15 (quinze) minutos antes a partir da inclinação dos valores da glicemia, indicando os níveis de glicose (estável, em aumento ou em diminuição). (SBD, 2019).

Tabela 01 – Interpretação das setas de tendência, advinda do SMCG

↑	O nível de glicose aumenta rapidamente (> 2 mg/dL por minuto) ou aumenta em 60 a 90 mg/dL em 30 minutos.
	O nível de glicose aumenta mais gradualmente (1 a 2 mg/dL por minuto) ou em 30 a 60 mg/dL em 30 minutos. → Glicose estável ou com alterações mínimas (< 1 mg/dL por minuto).
	Nível de glicose diminuindo gradualmente (1 a 2 mg/dL por minuto) ou em 30 a 60 mg/dL em 30 minutos.
↓	Nível de glicose diminuindo rapidamente (> 2 mg/dL por minuto) ou diminuído em 60 a 90 mg/dL em 30 minutos.

Fonte: Adaptada de Mohan V et al. (2016)

Vale ressaltar que o parâmetro de uma leitura atual de glicose junto a seta de tendência direcional torna-se importante para auxiliar em uma tomada de decisão no autogerenciamento da diabetes.

O uso das setas de tendência não indica nenhum sinal sonoro quanto ao nível proposto da futura glicemia. Se a seta de tendência estiver para baixo e chegar ao nível de hipoglicemia, o usuário não é alertado sobre o ocorrido. Isto ocorre porque o aparelho leitor não está vinculado a nenhum aplicativo ou dispositivo telefônico, sendo assim não fornece compartilhamento em tempo real.

3.2 APRENDIZADO DE MÁQUINA

Aprendizado de Máquina (AM) é uma subárea da Inteligência Artificial que diz respeito aos algoritmos capazes de aprender com os dados. O AM torna possível a análise e obtenção de novos padrões e associações para problemas que outrora

eram tratados apenas por pessoas especialistas em um determinado campo (JORDAN; MITCHELL, 2015).

O processo de aprendizado começa com exemplos para procurar padrões nos dados fornecidos e tomar melhores decisões no futuro com base nesses exemplos. O objetivo principal é permitir que os computadores aprendam automaticamente sem intervenção ou assistência humana e se ajustem de acordo com o surgimento de novos dados.

Esses algoritmos são treinados a partir de dados coletados e processados de acordo com o objetivo da tarefa. Dessa forma, é possível fazer previsões, como: classificação de imagens, previsão de preços, reconhecimento de padrões, sistemas de recomendação e etc. De acordo com Carvalho (2011), os algoritmos de AM geralmente seguem dois paradigmas: aprendizado supervisionado e aprendizado não supervisionado.

Os algoritmos de aprendizado supervisionado podem aplicar o que foi aprendido no passado a novos dados, usando exemplos rotulados para prever eventos futuros. A partir da análise de um conjunto de dados de treinamento conhecido, o algoritmo de aprendizado supervisionado produz uma função inferida para fazer previsões sobre os valores de saída. Esses algoritmos também podem comparar sua saída com a saída real (pretendida) e encontrar erros para modificar o modelo e, dessa forma, aumentar a performance dos indicadores de avaliação do modelo. Os algoritmos se diferem de acordo com o tipo dos rótulos dos dados, discreto ou contínuo, tendo assim, respectivamente, tarefas de classificação ou regressão.

Por outro lado, os algoritmos de aprendizado não supervisionados são usados quando as informações de treinamento não são classificadas e rotuladas. Com isso, os algoritmos levam os sistemas a uma função para descrever uma estrutura oculta a partir de dados não rotulados. O sistema não descobre a saída correta, mas explora os dados e pode fazer inferências de conjuntos de dados para descrever estruturas ocultas de dados não rotulados. Essas tarefas podem ser divididas em tarefas de agrupamento ou associação.

3.3 ALGORITMOS DE APRENDIZADO SUPERVISIONADO

O aprendizado supervisionado pode ser separado em dois tipos de problemas: classificação e regressão. A classificação usa um algoritmo para atribuir os dados de teste em categorias específicas. Ele reconhece entidades específicas dentro do conjunto de dados e tenta tirar algumas conclusões sobre como essas entidades devem ser rotuladas ou definidas. Os algoritmos de classificação mais comuns são: k-vizinho mais próximo (kNN), máquinas de vetor de suporte (SVM), árvores de decisão, classificadores de floresta aleatória.

A regressão é usada para entender a relação entre as variáveis dependentes e independentes. É comumente usada para fazer projeções, como determinar o preço de um produto em uma determinada região. Os algoritmos mais populares são: regressão linear, regressão logística e regressão polinomial.

Alguns algoritmos podem ser projetados especificamente para classificação, como regressão kNN, ou regressão, como regressão linear. Alguns podem ser usados para ambos os tipos de problemas com pequenas modificações, como Redes Neurais Artificiais (RNA).

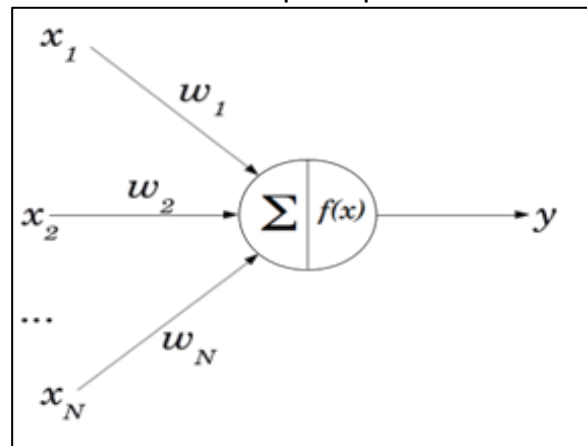
3.4 REDES NEURAIIS ARTIFICIAIS (RNA)

Uma RNA é um algoritmo projetado para assemelhar-se à forma que o cérebro humano analisa e toma uma decisão (SILVA, 2010). O cérebro humano tem bilhões de neurônios e cada neurônio é composto de um corpo celular responsável por computar informações, transportar para outros neurônios e fornecer uma saída final. Assim, as redes neurais podem ser classificadas em diferentes tipos, usados para diferentes propósitos.

3.4.1 Perceptron

A perceptron é uma rede neural que possui um único neurônio, chamado de nó ou unidade, e é a forma mais simples de uma RNA, onde o neurônio recebe as entradas x_n , multiplica cada entrada ao seu peso correspondente w_n , soma todos os resultados e passa essa soma por uma função não-linear, $f(x)$, chamada função de ativação, para produzir uma saída y .

Figura 04 – Representação gráfica de uma rede perceptron



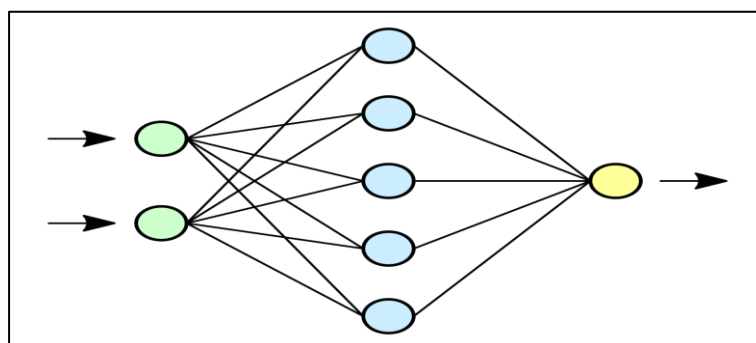
Fonte: Autoria própria

Essa rede é capaz de resolver problemas lineares, como classificações binárias. Uma função de ativação, geralmente, utilizada em uma rede perceptron é a função sigmóide. Essa função mapeia qualquer entrada numérica real para um valor que é 0 ou 1, $f(x) = \frac{1}{1+e^{-x}}$. A função de ativação de uma unidade tem como objetivo transformar a soma dos resultados das multiplicações em uma saída interpretável.

3.4.2 Perceptron Multi-camadas

Perceptron multi-camadas, ou *Multi-layer perceptron* (MLP), são compostas por camadas de entrada e saída e uma ou mais camadas ocultas com muitos neurônios empilhados juntos. E, enquanto na perceptron o neurônio tinha uma função de ativação que impõe um limiar, os neurônios em uma MLP podem usar qualquer função de ativação.

Figura 05 – Representação gráfica de uma rede perceptron multi-camadas



Fonte: (DAKE, 2006)

A figura acima representa a arquitetura de uma MLP com três camadas: camada de entrada, camada oculta e camada de saída. Cada unidade conecta-se a outra e tem um peso e um limiar associado. Se a saída de qualquer unidade individual estiver acima do valor limite especificado, essa unidade será ativada e enviará dados para a próxima camada da rede. Caso contrário, nenhum dado passará para a próxima camada da rede. Esse tipo de arquitetura chama-se de *feedforward*.

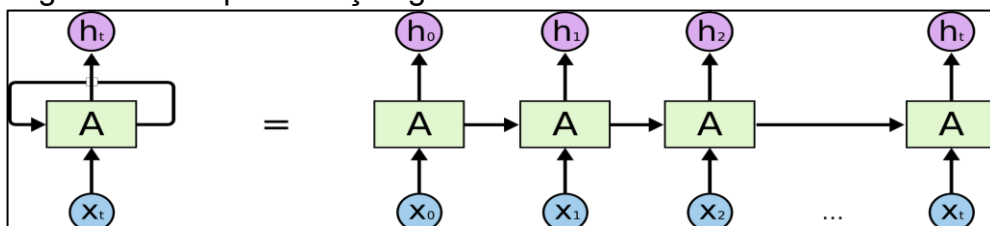
O treinamento dessa rede utiliza um algoritmo amplamente conhecido e utilizado nas RNAs. O *backpropagation* (RUMELHART, 1986), retropropagação, é um método que faz o ajuste dos pesos de uma rede neural com base na taxa de erro obtida na época anterior (ou seja, interação). O ajuste adequado dos pesos permite reduzir as taxas de erro e tornar o modelo preciso, aumentando a sua generalização e precisão.

Dessa forma, uma determinada quantidade de épocas é definida para o treinamento de uma rede neural. A escolha da quantidade de épocas para o treinamento de uma rede neural deve ser feita com cuidado e pode variar de acordo com o problema e número de amostras no conjunto de dados. Caso seja escolhida uma quantidade alta, o modelo pode perder a capacidade de generalização e ficar enviesado pelos dados passados no treinamento. Esse problema é chamado de *overfitting*.

3.4.3 Redes Neurais Recorrentes

As redes neurais recorrentes (RNNs) são identificadas por seus *loops* de *feedback*. Esses algoritmos de aprendizado são aproveitados, principalmente, ao usar dados de séries temporais para fazer previsões sobre resultados futuros, como previsões do mercado de ações ou previsão de vendas.

Figura 06– Representação gráfica de uma rede neural recorrente



Fonte: (CHEN et al, 2016)

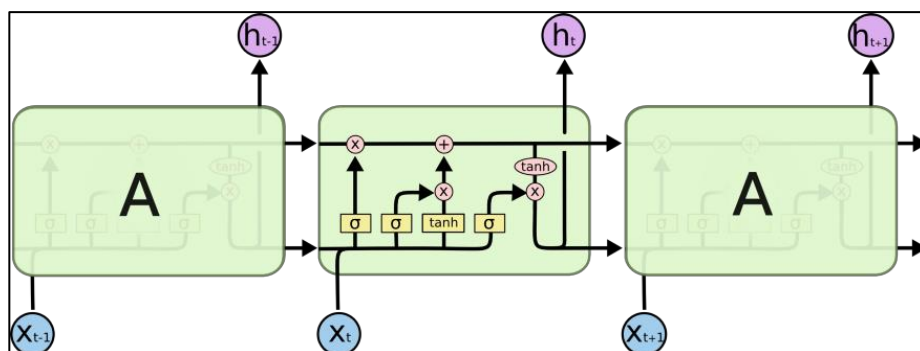
A cada passo de tempo t , a rede se expande para obter a saída no passo de tempo $t + 1$. Cada rede é semelhante à rede neural perceptron. Assim, a rede calcula os valores das unidades ocultas e a saída após os intervalos de tempo. Cada camada recorrente tem dois conjuntos de pesos: um para a entrada e outro para a unidade oculta. A última camada calcula a saída final para o passo de tempo t .

Existe um problema comum em grandes redes recorrentes: o desaparecimento do gradiente descendente. Esse problema acontece devido ao algoritmo de treinamento da RNN que fica perdido ao tentar calcular novos pesos para entradas muito longas e, assim, causando o esquecimento do que foi aprendido no longo prazo.

Uma solução para esse problema foi proposta por Hochreiter e Schmidhuber (1997). Eles criaram uma rede recorrente chamada LSTM (*Long-Short Term Memory*, Memória de Longo Prazo) que é treinada usando o algoritmo *Backpropagation Through Time* e supera o problema do gradiente descendente. A rede LSTM pode ser usada para criar grandes redes recorrentes para resolver problemas de sequência difíceis em aprendizado de máquina e obter bons resultados.

As redes LSTM possuem blocos de memória conectados por meio de camadas. Cada bloco contém portas que gerenciam o estado e a saída. Essas portas tornam a rede mais inteligente que um neurônio comum e adicionam uma memória para sequências recentes. Esses blocos operam sobre uma sequência de entrada e cada porta dentro de um bloco usa as unidades de ativação sigmoidal para controlar se elas serão acionadas ou não, condicionando a mudança de estado e a adição de informações que fluem pelo bloco.

Figura 07 – Arquitetura de uma rede LSTM



Fonte: CHEN, 2016

Existem três tipos de portas dentro de uma unidade: a *forget gate*, decide condicionalmente quais informações ela exclui do bloco; a *input gate*, decide quais valores de entrada podem atualizar o estado da memória; e a *output gate*, que decide o que produzir com base na entrada e na memória do bloco. As portas de cada unidade possuem pesos que são aprendidos durante o procedimento de treinamento.

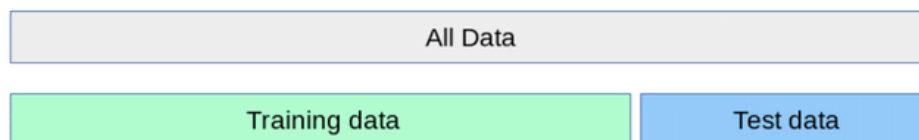
3.5 VALIDAÇÃO CRUZADA

Um bom modelo de aprendizado de máquina precisa ser o mais genérico possível para que esse algoritmo possa ser eficaz em várias amostras de entrada. Ou seja, o modelo não sofre degradação de desempenho em novas amostras da mesma distribuição dos dados de treinamento. A validação cruzada é um método para avaliar um modelo de aprendizado de máquina e testar o seu desempenho.

3.5.1 Holdout

O método *Holdout* é o tipo mais simples de validação cruzada. A base de dados é separada em duas bases, chamadas de base de treinamento e base de teste. As amostras são divididas aleatoriamente seguindo uma determinada proporção (geralmente, 70% para a base de treino e 30% para a base de teste).

Figura 08 – Método *Holdout*



Fonte: SCUDILIO, 2020.

O modelo é treinado utilizando apenas a base de treinamento. Após o treinamento, o modelo faz a previsão dos valores de saída para os dados na base de teste (o modelo nunca viu esses valores de saída antes). A vantagem deste método é que geralmente não leva muito tempo para calcular. No entanto, sua avaliação pode ter uma variação. A avaliação pode depender muito de quais amostras estavam presentes na base de treinamento e quais estavam na base de

teste e, portanto, a avaliação pode ser significativamente diferente dependendo de como a divisão é feita.

3.5.2 *k-Fold*

O método *k-Fold* é uma forma de melhorar o método *Holdout*. A base de dados é dividida em k subconjuntos e o método *Holdout* é repetido k vezes. A cada vez, um dos k subconjuntos é usado como base de teste e os outros $k-1$ subconjuntos são selecionados juntos para formar uma base de treinamento. O erro médio em todas as k tentativas é então calculado. (LOCA, 2020)

A vantagem desse método é que não importa muito como os dados são divididos. Cada amostra de dados fica em uma base de teste uma vez e fica em uma base de treinamento $k-1$ vezes. A variância do resultado diminui à medida que k aumenta.

Figura 09 – Método *k-Fold*



Fonte: (FEDOTENKOVA , 2016)

A desvantagem deste método é que o algoritmo de treinamento deve ser executado k vezes, o que significa que k vezes mais cálculos são necessários para realizar uma avaliação.

Esse método de análise de predição possibilita, portanto, avaliar a qualidade dos dados em relação à generalização do modelo. Quanto melhor for a avaliação do modelo no *k-Fold*, melhor ele performa para novas amostras que possam entrar.

4 MATERIAIS E MÉTODOS

4.1 CONSIDERAÇÕES ÉTICAS

Este estudo seguiu as determinações da Resolução 466/2012 do Conselho Nacional de Saúde, que legisla sobre diretrizes e normas regulamentadoras de pesquisa envolvendo seres humanos (BRASIL, 2012). O projeto de pesquisa foi submetido ao Comitê de Ética em Pesquisa da Universidade Brasil e da Instituição em estudo.

O Parecer Consubstanciado Nº 4.506.553 (em Apêndice G) da instituição proponente Universidade Brasil aprovou a presente pesquisa para que assim fosse realizada com legalidade e ética, garantindo carácter científico e acadêmico.

Os participantes foram convidados a fazer parte do estudo e esclarecidos acerca dos objetivos e finalidades da pesquisa, bem como dos procedimentos para coleta dos dados. Aos que aceitaram participar, foram entregues os questionários (em Apêndice A) sobre a caracterização dos participantes da pesquisa e o mapeamento da situação de saúde do participante (em Apêndice B), o aplicativo de contagem de carboidratos (para que preencham o consumo de carboidrato diário) e o Termo de Consentimento Livre e Esclarecido – TCLE (em Apêndice C) em duas vias, permanecendo uma com o participante. No TCLE constam as informações sobre a pesquisa, telefone e e-mail do pesquisador responsável para contato e caso houvesse necessidade de algum esclarecimento pertinente ao estudo.

Dúvidas que surgiram no transcurso da pesquisa foram sanadas, como questões relacionadas ao anonimato dos dados e à desistência ao longo da pesquisa. Diante disso, foi garantida a participação voluntária e a exclusão do consentimento a qualquer tempo, sem prejuízos ou constrangimento, além de garantir o anonimato das informações pessoais e a privacidade do participante.

Cabe destacar que a realização do estudo não acarretou nenhum dano ao participante e que o pesquisador esteve sempre à disposição para esclarecimentos e orientações relacionadas ao tema, que pudessem surgir durante a coleta dos dados.

4.2 TIPO DE ESTUDO: DESENHO

O presente trabalho configura-se em uma pesquisa qualitativa e quantitativa, partindo da perspectiva de que o trabalho busca a partir de dados quantitativos uma medição da glicemia com uso de uma tecnologia artificial, ou seja, sugerindo uma proposta de intervenção a partir de uma pesquisa qualitativa. Trata-se de um estudo transversal definido por uma população de interesse, pois é uma análise que ocorre em um período de tempo pré-definido (de 14 em 14 dias) dos dados de nível glicêmico.

Figura 10 – O tamanho do sensor do freestyle libre x moeda de 1 centavo americano



Fonte: (ABBOT, 2020)

O sensor tem o tamanho e forma de uma pequena moeda que fica aderente à epiderme. Comercialmente é veiculado como *FreeStyle Libre®* da Abbott, composto por um sensor, um leitor e um software. (HAAK et al, 2017; CARDOSO et al, 2018). Tem-se como amostra da presente pesquisa portadores e não portadores de diabetes mellitus tipo II. O estudo é analítico por realizar comparação de grupo de estudo com grupo controle em um retrato de sua situação atual.

4.3 LOCAL DE COLETA DE DADOS

A pesquisa foi originada e orientada pelo professor Dr. Marcello Magri Amaral do Instituto Científico e Tecnológico da Universidade Brasil, campus Itaquera. Coletaram-se os dados obtidos dos sujeitos participantes em São Luís – MA, em

colaboração com a Universidade Estadual do Maranhão, campus Paulo VI, pelo Posto de Serviço Social e Médico da comunidade acadêmica, onde já havia sido coordenador. Isso ajudou no acesso aos voluntários, já que a maioria deles eram do próprio campus da UEMA de São Luis, cidade onde o pesquisador reside e trabalha.

4.4 PARTICIPANTES

Para realização deste estudo foram selecionados 10 (dez) participantes portadores de diabetes mellitus tipo II e 10 (dez) participantes não portadores de DM2, entre a faixa etária de 30 a 65 anos, sendo 2 (dois) obesos, 15 (quinze) sobrepesos e 3 (três) com peso normal.

Os sujeitos da pesquisa são naturais da cidade de São Luís – MA, que seguem os critérios de inclusão e exclusão para os voluntários com DM2 descritos abaixo:

- a) **Critérios de inclusão** – Portadores de Diabetes Mellitus tipo II, na faixa etária entre 30 e 65 anos.
- b) **Critérios de exclusão** – Portadores de Diabetes mellitus tipo I e gestacional, faixa etária entre 0 e 29 anos e doze meses e acima de 66 anos.

Os critérios de inclusão e exclusão para os voluntários sem DM2 foram:

- a) **Critérios de inclusão** – Pessoas que se reconhecem sem Diabetes Mellitus tipo I e II, na faixa etária entre 30 e 65 anos.
- b) **Critérios de exclusão** – Portadores de Diabetes mellitus tipo I, II e gestacional, faixa etária entre 0 e 29 anos e doze meses e acima de 66 anos.

Os voluntários foram antecipadamente submetidos a exame físico, aferição de pressão arterial, aferição de glicemia, avaliação antropométrica, assinatura e entrega ao pesquisador do termo de consentimento livre e esclarecido (TCLE), questionário sociodemográfico e hábitos de vida, contagem de calorias através de questionário que incluíram respostas sobre café da manhã, almoço, lanche, jantar e refeição extra.

Vale lembrar que o sensor fora posicionado no braço direito dos voluntários pelo pesquisador responsável em dias distintos devido à inviabilidade da obtenção dos sensores em decorrência do custo que o aparelho proporciona e também por alguns apresentarem problemas no transcurso da pesquisa, necessitando à espera

da troca do aparelho. A retirada do aparelho se deu também pelo pesquisador responsável ao final do período da medição do sensor.

Para que a pesquisa ocorresse de forma eficaz, o voluntário recebeu um dispositivo leitor para medição do nível de glicose onde pode ser aferido pelo indivíduo ao aproximar o dispositivo leitor do sensor colado à pele.

A medição pode ser realizada no mínimo 3 vezes por dia pelo próprio voluntário, para que os dados armazenados na memória do sensor sejam transferidos para o dispositivo leitor e, assim, armazenar toda a série temporal dos dados.

4.4.1 Caracterização sociodemográfica e de hábitos de vida

Logo após a assinatura do TCLE, o participante voluntário preencheu o questionário sociodemográfico e de hábitos de vida (Apêndice A e B respectivamente), que foram utilizados para estabelecimento de possíveis correlações relevantes exploradas no desenvolvimento deste projeto.

Além dos questionários, o aplicativo de contagem de carboidrato da Sociedade Brasileira de Diabetes permitiu determinar a porcentagem de macronutrientes, como: carboidratos, proteínas e gordura; e encontra-se disponível gratuitamente para IOS e Android. A escolha desse aplicativo tornou-se viável por ser mais atual e utilizado pela Sociedade Brasileira de Diabetes associado aos dados coletados. Compreende-se com isso que os hábitos de vida podem estar correlacionados com a contagem de carboidratos e calorias. Esse fato pode ser melhor visualizado quando os participantes detalham suas principais refeições: café da manhã, almoço, lanche, jantar e refeições extras.

4.5 INSTRUMENTOS UTILIZADOS PARA COLETA DE DADOS

4.5.1 Avaliação Física

Em todos os voluntários foram aplicadas uma anamnese, investigando os principais sinais vitais: pressão arterial (PA), temperatura corporal e glicemia capilar. Essa avaliação física torna-se imprescindível para se investigar os parâmetros iniciais de diagnóstico para qualquer tipo de doença.

Segundo a OMS, a pressão arterial normal caracteriza-se quando as artérias não se encontram sob estresse e não há sobrecarga do coração. Atualmente, os níveis de classificação arterial para adolescentes, adultos e idosos assim se classificam:

- PRESSÃO ARTERIAL NORMAL – pressão sistólica menor que 120 mmHg e pressão diastólica menor que 80 mmHg.
- PRÉ-HIPERTENSÃO – pressão sistólica entre 120 e 129 mmHg ou pressão diastólica menor que 80 mmHg.
- HIPERTENSÃO ESTÁGIO 1 – pressão sistólica entre 130 e 139 mmHg ou pressão diastólica entre 80 e 89 mmHg.
- HIPERTENSÃO ESTÁGIO 2 – pressão sistólica acima de 140 mmHg ou pressão diastólica acima de 90 mmHg.
- CRISE HIPERTENSIVA – pressão sistólica acima de 180 mmHg ou pressão diastólica acima de 110 mmHg.

Tabela 02 – Classificação da pressão arterial em adultos

CRITÉRIOS DIAGNÓSTICOS E CLASSIFICAÇÃO	PAS (mmHg)	PAD (mmHg)
Normal	< 120	< 80
Pré-hipertensão	120-139	80-89
Hipertensão Estágio 1	140-159	90-99
Hipertensão Estágio 2	>160	>100

Fonte: Autoria Própria

Para fins de melhor entendimento, de acordo com a tabela 02, o estágio do quadro hipertensivo deve-se ao valor mais alto de sistólica ou diastólica. E deve-se utilizar a maior para a classificação de estágio em pacientes quando as pressões sistólica e diastólica situarem-se em estágios diferentes.

A aferição da pressão arterial foi realizada com o aparelho denominado de Esfigmomanômetro, de marca *Caretech*; e com estetoscópio, de marca *Premium*, para a ausculta dos batimentos cardíacos. Para a temperatura corporal, foi utilizado o termômetro de marca *Termomed*. E para a aferição de glicemia capilar, fez-se uso de aparelho da marca *Accu-Chek*.

4.5.2 Avaliação Antropométrica

Essa avaliação visa avaliar as dimensões físicas de uma pessoa. Nessa pesquisa primou-se pelo peso, altura e o IMC (Índice de Massa Corporal) do voluntário.

A aferição do peso do voluntário foi realizada por balança mecânica de controle corporal, de marca *G-Tech*, capacidade máxima de 150kg. Conforme as orientações da OMS-1995, orientou-se que os participantes vestissem roupas leves e subissem de costas, distribuindo o peso sobre o aparelho, para a realização da antropométrica.

Foi utilizado trena antropométrica, de marca *Cescorfe* com 2 metros de comprimento e 6 centímetros de largura. Conforme orientações da OMS-1995, voluntário ficou em posição ereta de costa para uma parede, mantendo a visão em linha perpendicular ao corpo para a realização do aferimento.

Para avaliar se o peso está adequado à altura de uma pessoa, esta pesquisa fez uso do indicador do IMC (Índice de Massa Corporal). A OMS orienta a utilização desse indicador de massa corpórea, por ser obtido de forma rápida e sem nenhum custo adicional, apenas fazendo a medição do peso e a altura. Para se obter o resultado do IMC, dividiu-se o peso pela altura elevado ao quadrado, representado pela fórmula a seguir: $P(Kg)/h^2(m)$.

Por meio deste indicador, pode-se inferir a possíveis problemas de saúde, que serão verificados através de exames solicitados por um médico. A OMS sugere a seguinte classificação para os resultados do IMC:

Tabela 03 – Classificação do IMC de acordo com a OMS

IMC	Classificação
Abaixo de 18,5	Baixo peso
Entre 18,6 e 24,9	Peso normal
Entre 25 e 29,9	Sobrepeso
Entre 30 e 34,9	Obesidade grau I
Entre 35 e 39,9	Obesidade grau II
Acima de 40	Obesidade grau III

Fonte: Autoria própria

De acordo com a Tabela 03, os resultados do IMC indicam se o indivíduo está com *baixo peso*, *peso normal*, *sobre peso*, *obesidade grau I*, *obesidade grau II* e

obesidade grau III. A importância dessa medição para a presente pesquisa deve-se à correlação que existe sobre a gordura corporal com a alteração da glicemia. Essa variável é estudada aqui para identificar um dos pontos que levam a diabetes em um indivíduo.

4.5.3 Contagem de Carboidrato

A Contagem de Carboidratos ajuda a melhorar o controle metabólico da pessoa com diabetes, baseando-se no fato de que 100% dos carboidratos ingeridos transformam-se em glicose, assim como 50% das proteínas e 10% dos lipídios.

As Associações Britânica e Americana de Diabetes, a partir de 1980, passou a controlar a diabetes com uma dieta controlada em gorduras, rica em carboidratos complexos e fibras alimentares, respeitando a individualização do plano alimentar. Cabe aqui lembrar que “os carboidratos são os maiores responsáveis pela glicemia pós-prandial, uma vez que são praticamente todos convertidos em glicose”. Por isso, a necessidade da contagem de carboidrato consumida por refeição, pois a quantidade a ser ingerida deverá obedecer aos padrões diários estabelecidos.

Diante disso, a contagem de carboidratos foi realizada em um período de 14 dias em todas as refeições diárias (café da manhã, lanches, almoço e ceia) concomitantemente à aplicação do sensor CGM. Lembrando que esse procedimento ocorre por meio de um aplicativo desenvolvido pela Sociedade Brasileira de Diabetes – SBD, a fim de melhorar a qualidade de vida dos portadores de diabetes mellitus.

O *Manual da contagem de carboidratos* desenvolvido pela SBD (Sociedade Brasileira de Diabetes) mostra quais os alimentos são considerados na contagem.

4.6 METODOLOGIA DE ANÁLISE DOS DADOS

Os dados do dispositivo do leitor foram descarregados em computador via porta USB, resguardadas as informações pessoais dos voluntários e utilizados métodos de análise de séries temporais a estes dados, a fim de se extrair parâmetros que levem à predição do nível glicêmico.

As informações do aplicativo *Contagem de carboidratos*, da Sociedade Brasileira de Diabético, e dos questionários sociodemográficos e de hábitos de vida foram tabelados e estatisticamente analisados para caracterização dos participantes e de sua situação de saúde. E, por fim, explica-se a criação do modelo LSTM por meio do *Manual Search*, método de inteligência artificial escolhido, para estabelecimento de uma técnica de predição do nível glicêmico.

4.6.1 Processamento dos dados

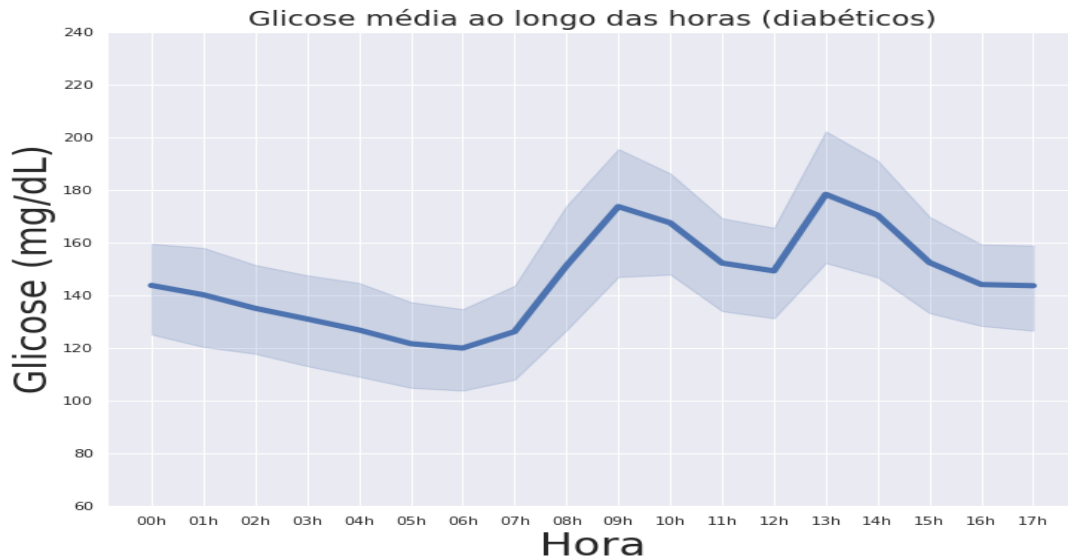
A previsão do valor da glicose ao longo do tempo é uma tarefa de regressão que apresenta um bom desempenho por um modelo supervisionado. Todo o trabalho foi desenvolvido na linguagem de programação Python na plataforma Jupyter (KLUYVER, 2016).

Os valores da glicose (mg/dL) foram coletados a cada 15 minutos durante 24 horas, totalizando 96 observações, levando em consideração que 1 hora tem 4 observações, tem-se a seguinte equação: $n^{\circ} \text{ observações} = n^{\circ} \text{ horas} \times 4$. Assim, a cada 15 minutos obteve-se o valor referente ao nível da glicose e, ao longo de um dia completo, têm-se 96 observações.

Chegou-se analisar outras durações que confirmaram que aumentar demais o tamanho dessa série temporal não resultaria em uma melhoria na qualidade da predição. E diminuir esse tempo acarretaria uma quantidade menor do tempo predito. Portanto, tomou-se como um intervalo razoável 12 horas de dados contínuo de glicose, para se ter um valor de 5 horas preditas.

De acordo com o objetivo deste trabalho, ao verificar se é possível a predição da glicose dado valores passados, utilizou-se 12 horas de dados para tentar prever 5 horas a frente. Convertendo esses valores para número de observações, têm-se, respectivamente, 48 e 20 observações, totalizando 68 observações. Ressalta-se que para cada 1 hora tem-se 4 observações.

Figura 11 – Glicose média ao longo das observações dos dados dos voluntários diabéticos

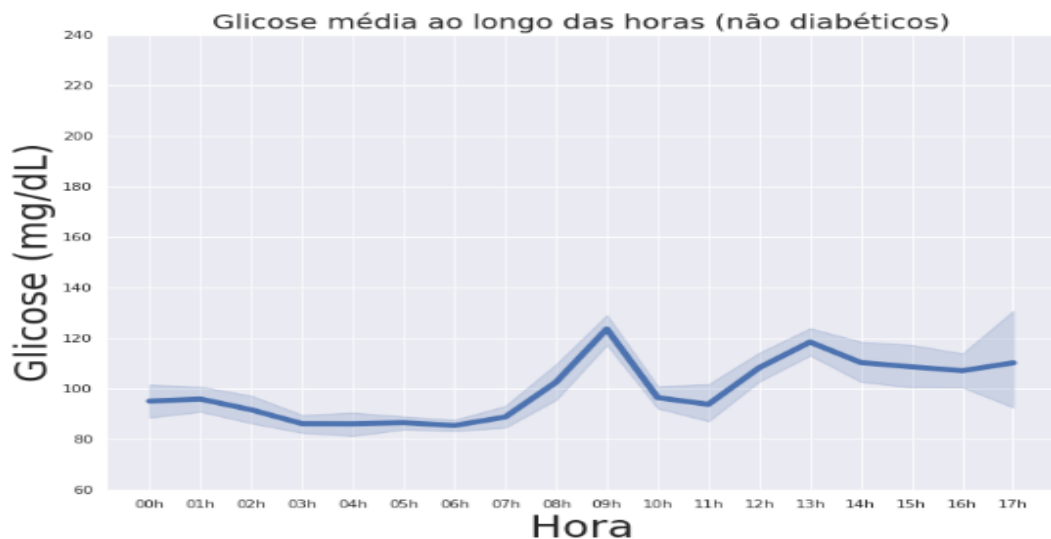


Fonte: Autoria própria

De acordo com Figura 11, há um pico de glicemia de 170 mg/dL e 180 mg/dL entre os participantes diabéticos, correspondentes às 9h e 13h, momento em que ocorre as primeiras refeições diárias, coincidindo com o tempo em que a ingestão de macronutrientes (carboidratos, frutose, lactose e outros) leva para afetar a glicemia no organismo do indivíduo.

No que se refere à glicose no grupo de não diabéticos, no Figura 12 observa-se um pico de glicemia de 130 mg/dL e 120 mg/dL correspondentes às 9h e 13h. Além disso, apresenta uma frequência glicêmica de 110 mg/dL.

Figura 12 – Glicose média ao longo das observações dos dados dos voluntários não diabéticos



Fonte: Autoria própria

Ao todo obteve-se 275 amostras, mas devido a alguns problemas durante a coleta (como dados faltantes ao longo do dia, amostras continham menos que 68 observações), algumas não foram consideradas para o treinamento do modelo. Constaram, dessa forma, 216 amostras selecionadas.

É possível observar, no Figura 11, que a média e a variabilidade do comportamento glicêmico dos voluntários diabéticos ao longo das 12 horas são maiores que dos voluntários não diabéticos, ocorrendo entre 120 mg/dL e 180 mg/dL. Nota-se também que os picos no valor da glicose são maiores entre os diabéticos, às 9 horas e às 13 horas do dia.

4.7 CRIAÇÃO E TREINAMENTO DO MODELO LSTM

Para este trabalho foi utilizado uma rede neural recorrente LSTM para a prever o valor da glicose dado uma série de valores passados. Dadas 48 observações³, equivalentes a 12 horas, teve como objetivo gerar mais 20 observações, equivalente a 5 horas, com base nas observações anteriores. Isso significa que a cada 12 horas de dados coletados pode-se prever 5 horas do comportamento glicêmico dos participantes.

Dessa forma, uma determinada quantidade de épocas é definida para o treinamento de uma rede neural, bem como outros hiperparâmetros, tais como função de perda, otimizador, função de ativação, número de camadas ocultas, número de neurônios etc (BERGSTRÄ; BENGIO, 2012). A escolha de cada hiperparâmetro para o treinamento de uma rede neural deve ser feita com cuidado e pode variar de acordo com o problema e número de amostras no conjunto de dados.

A escolha da quantidade de camadas ocultas em uma RNA e a quantidade de neurônios contidos em cada camada não seguem nenhum padrão universal. Entretanto, o cuidado que se deve ter durante a arquitetura de um modelo é com o *Underfitting* e o *Overfitting*. *Underfitting* é quando o modelo não consegue aprender com os dados passados no treinamento; isso acontece, por exemplo, quando a rede tem apenas uma camada. De forma contrária, o *Overfitting* é quando o modelo perde a capacidade de generalização e fica enviesado pelos dados passados no

³ As observações aqui citadas referem-se à glicemia aferida a cada 15 minutos.

treinamento, o que pode acontecer caso seja adicionado um número épocas alto ou várias camadas ocultas.

Utilizou-se as bibliotecas TensorFlow (2022) e Scikit-learn (PEDREGOSA, 2011) para a criação e avaliação do modelo LSTM. Fez-se a variação dos seguintes hiperparâmetros: número de camadas ocultas, número de neurônios em cada camada, número de épocas, função de ativação, otimizador e função de perda. O objetivo dessa variação é encontrar aqueles que dão melhores resultados na performance do modelo LSTM. O método escolhido para essa tarefa foi o Manual Search (BERGSTRA; BENGIO, 2012). Após isso, foram escolhidos os seguintes hiperparâmetros para o modelo: 70 épocas, função de ativação tangente hiperbólica, otimizador Adam, função de perda, sendo o erro quadrático médio (MSE). Então, criou-se um modelo com 5 camadas.

Para a avaliação do modelo, utilizou-se as seguintes métricas para modelos de regressão:

- MAE (*Mean Absolute Error*, ou erro absoluto médio) – representa a média da diferença absoluta entre os valores reais e previstos no conjunto de dados. Mede a média dos resíduos no conjunto de dados.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}|, \text{ onde } \hat{y} \text{ é o valor predito de } y.$$

- MSE (*Mean Squared Error*, ou erro quadrático médio) – representa a diferença quadrática média entre os valores reais e previsto no conjunto de dados. Mede a variância dos resíduos.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2, \text{ onde } \hat{y} \text{ é o valor predito de } y.$$

- RMSE (*Root Mean Square Error*, ou raiz do erro quadrático médio) – nessa métrica o erro retorna à unidade de medida do modelo. Assim como o MSE, ela mede a variância dos resíduos.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}, \text{ onde } \hat{y} \text{ é o valor predito de } y.$$

As 216 amostras de dados foram utilizadas na validação cruzada usando a técnica *k-Fold* com $k = 5$, ou seja, 5 pastas. Após essa validação, as 216 amostras foram divididas aleatoriamente em bases de treino e teste, numa proporção 70% e 30% respectivamente. Essa técnica de divisão é chamada de *holdout*.

Ademais, os dados foram normalizados em um intervalo de 0 a 1 antes de aplicar no modelo, utilizando o método *MinMaxScaler* da Scikit-learn. Essa

normalização faz com que a computação do ajuste dos pesos da rede neural fique menos robusta e complexa. Após a realização da predição, é necessário inverter o processo de normalização.

4.8 MODELO PERSONALIZADO

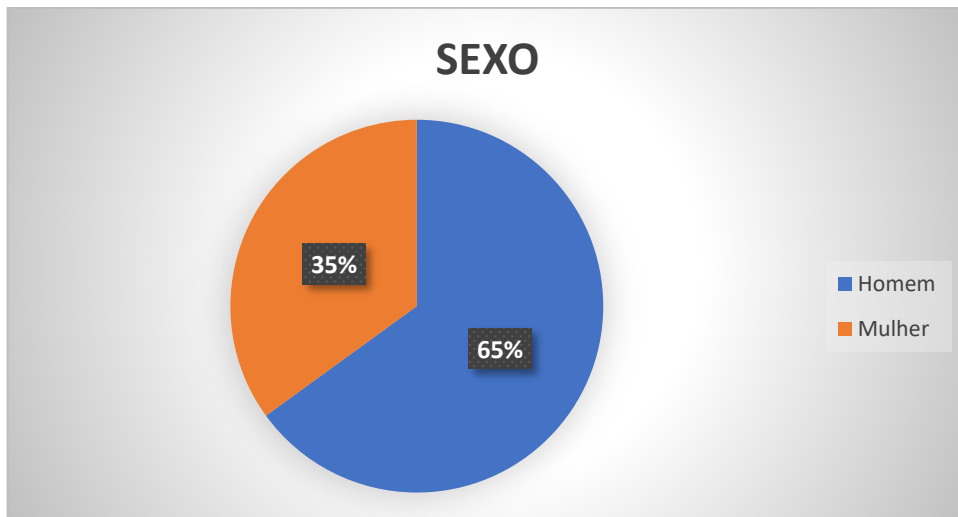
Selecionou-se os dados de dois voluntários: um diabético e um não diabético, para a criação de dois modelos LSTM individuais. Cada voluntário contém dados de 8 amostras com 68 observações. Dividiu-se os dados em uma proporção de 80% para a base de treinamento e 20% para a base de teste.

Foram escolhidos os seguintes parâmetros para cada um dos dois modelos LSTM: 70 épocas, função de ativação tangente hiperbólica, otimizador Adam, função de perda sendo o erro quadrático médio (MSE). Criou-se 4 camadas: adicionou-se 48 unidades na primeira camada, 200 unidades na segunda camada, 200 unidades na terceira camada e 20 unidades na quarta camada. A avaliação do modelo foi feita utilizando as mesmas métricas definidas acima.

5 RESULTADOS

A pesquisa contou com um total de 20 participantes, sendo 65% do sexo masculino e 35% do sexo feminino, de acordo com a Figura 13.

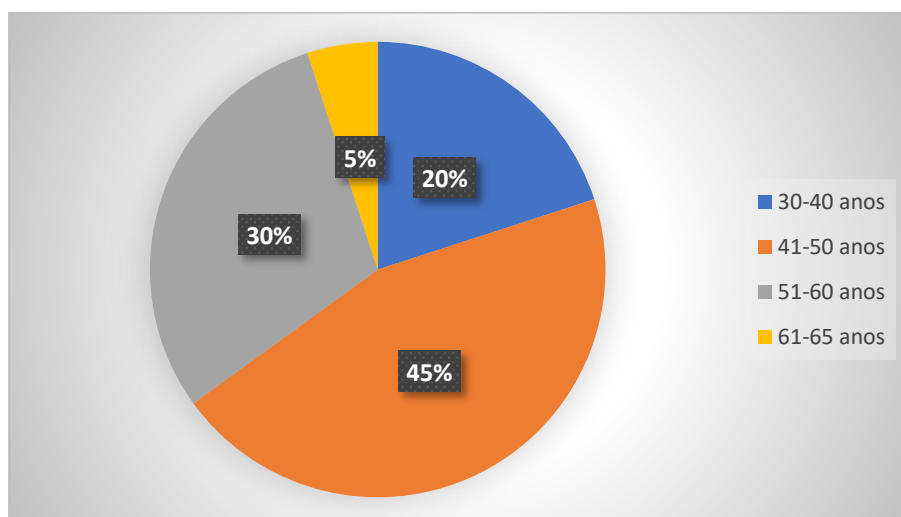
Figura 13 – Sexo dos voluntários



Fonte: Autoria própria

É importante lembrar que 50% dos voluntários são diabéticos e 50% não diabéticos. Dos sujeitos investigados, pode-se dizer que 20% estão na faixa etária de 30-40 anos, 45% com 41-50 anos, 30% com 51-60 anos e 5% com 61-65 anos, em Figura 14. A situação conjugal configura-se em 70% casados e 30% divorciados.

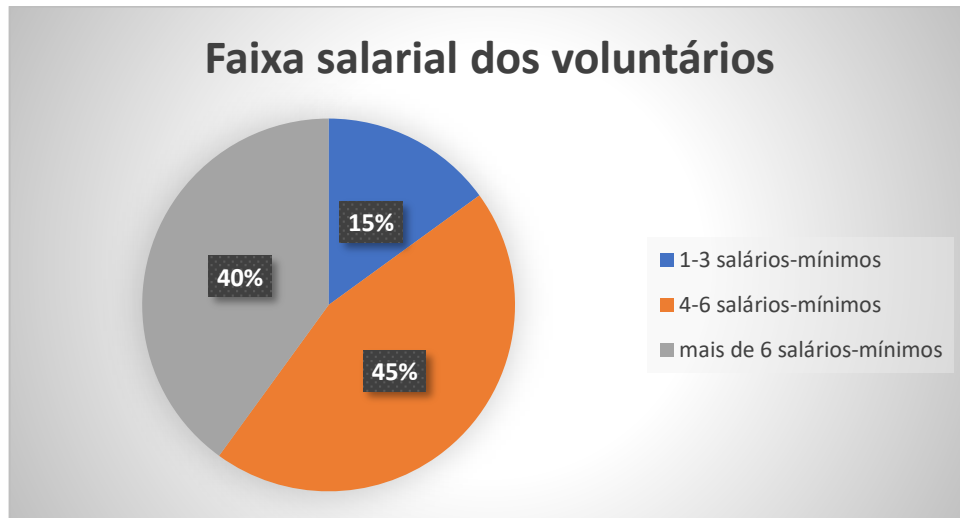
Figura 14 – Faixa etária dos voluntários



Fonte: Autoria própria

A renda familiar dos participantes varia com 15% dos participantes com 1-3 salários-mínimos, 45% com 4-6 salários e 40% com mais de seis salários (Figura 15).

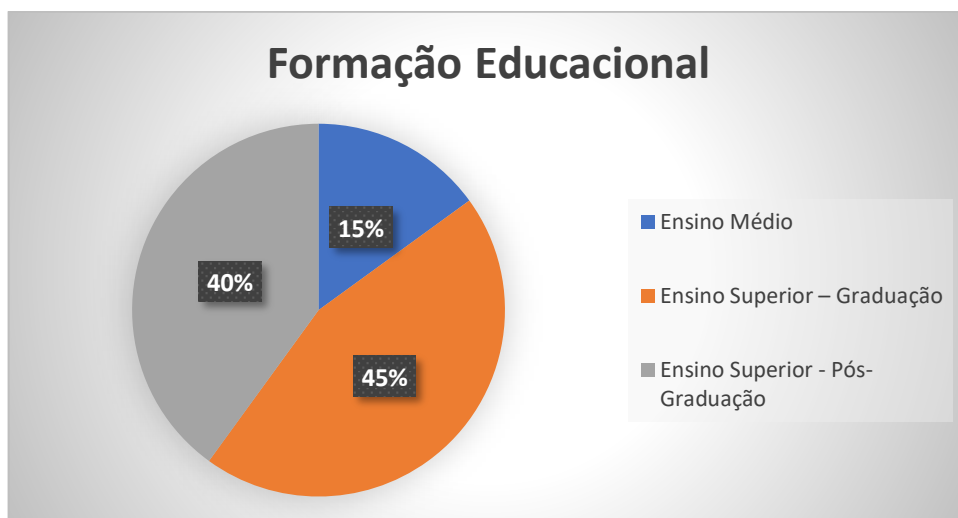
Figura 15 – Faixa salarial dos voluntários



Fonte: Autoria própria

A variação dessa renda condiz com o nível educacional dos sujeitos, em que apenas 15% possuem ensino médio, 45% ensino superior e 40% pós-graduação (Figura 16).

Figura 16 – Formação educacional salarial dos voluntários DB

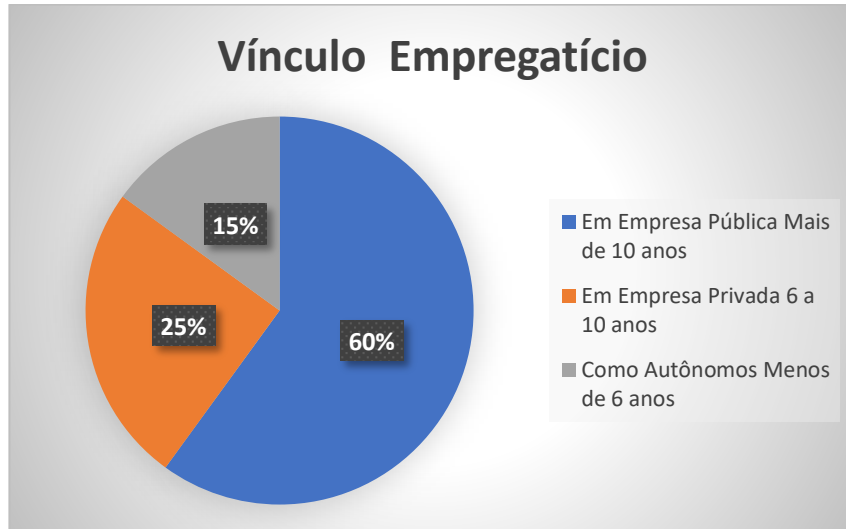


Fonte: Autoria Própria

Observou-se também que 60% trabalham em empresa pública, 25% em empresa privada e 15% são autônomos, sendo que, respectivamente, os primeiros contam com mais de 10 anos de trabalho, os segundos entre 6 e 10 anos e os

terceiros com menos de 6 anos (em Figura 17), com uma carga horária de no mínimo 40 horas semanais, com a maioria trabalhando em dois turnos e apenas duas pessoas no turno noturno.

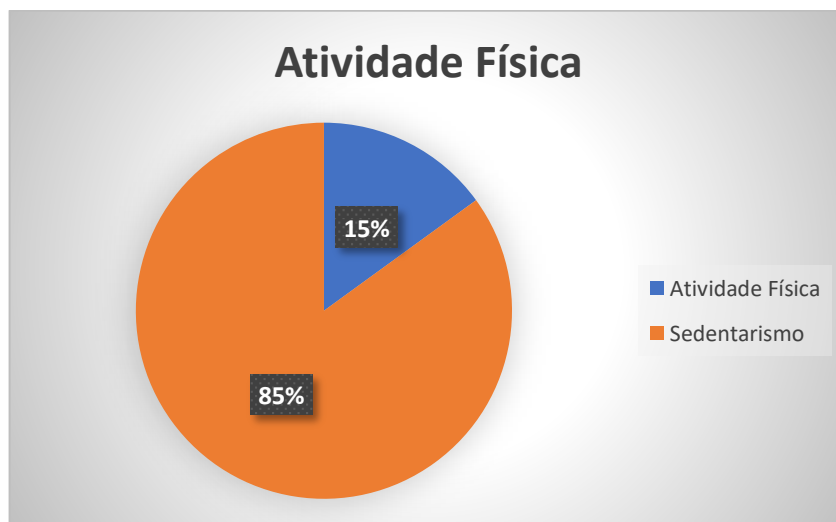
Figura 17 – Formação educacional salarial dos voluntários NDB



Fonte: Autoria própria

Quanto à situação de saúde do participante, apenas 15% fazem atividade física (caminhada) e 85% são sedentários (em Figura 18). Todos dispõem de algum tipo de lazer em seu tempo livre, como shopping, barzinhos, viagens, praias e outros.

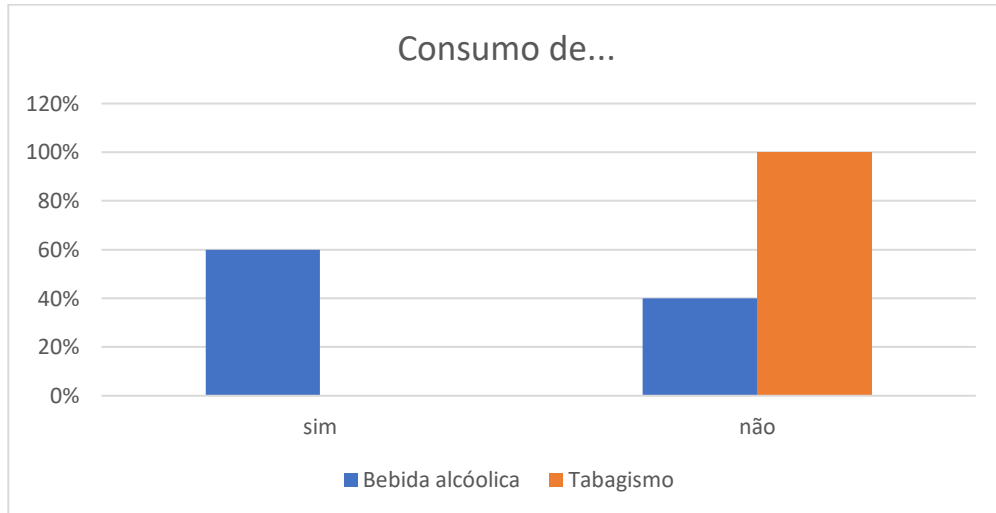
Figura 18 – Atividades dos voluntários



Fonte: Autoria própria

Das pessoas que têm diabetes, 30% fazem ingestão de bebidas alcóolicas e, dentre as pessoas não diabéticas, apenas 30% fazem uso de bebidas alcóolicas (Figura 19). E de todos os voluntários nenhum faz uso do consumo de tabaco.

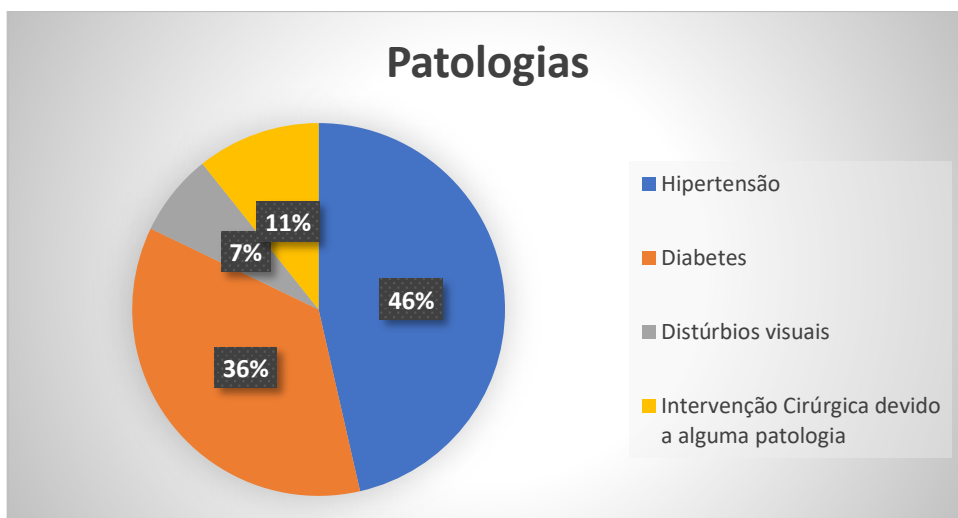
Figura 19 – Consumo de bebida alcóolica e tabagismo dos voluntários



Fonte: Autoria própria

No que se refere a patologias, 65% têm hipertensão, 50% diabetes, 10% distúrbios visuais e apenas 15% dos voluntários passaram por algum tipo de cirurgia (em Figura 20).

Figura 20 – Patologias e Intervenções Cirúrgicas dos voluntários



Fonte: Autoria própria

Vale a pena lembrar que uma patologia não anula a outra, por isso a presença de várias comorbidades descritas na Figura 20 em um mesmo voluntário.

Tabela 04 – Avaliação Física dos sinais vitais dos voluntários diabéticos

VOLUNTÁRIOS (Com Diabetes)	PA	TEMPERATURA	GLICEMIA
V01	133x98 mmHg	36,7°C	112 mg/dL
V02	112x85 mmHg	35,9°C	123 mg/dL
V03	141x86 mmHg	36,6°C	126 mg/dL
V04	114x101 mmHg	36,9°C	143 mg/dL
V05	114x110 mmHg	36,9°C	121 mg/dL
V06	115x12 mmHg	36,7°C	129 mg/dL
V07	133x110 mmHg	36,9°C	148 mg/dL
V08	113x110 mmHg	37,0°C	134 mg/dL
V09	112x80 mmHg	36,8°C	131 mg/dL
V10	113x90 mmHg	36,9°C	142 mg/dL

Fonte: Autoria própria

De acordo com a Tabela 04, dos voluntários com diabetes apenas 10% encontram-se com a pressão arterial estável, enquanto 90% apresentam a pressão arterial acima. No que se refere a temperatura, 100% mostram possuir em estado normal. Mas o fator glicêmico apresenta-se alterado bem acima do valor de referência que estabelece a normalidade.

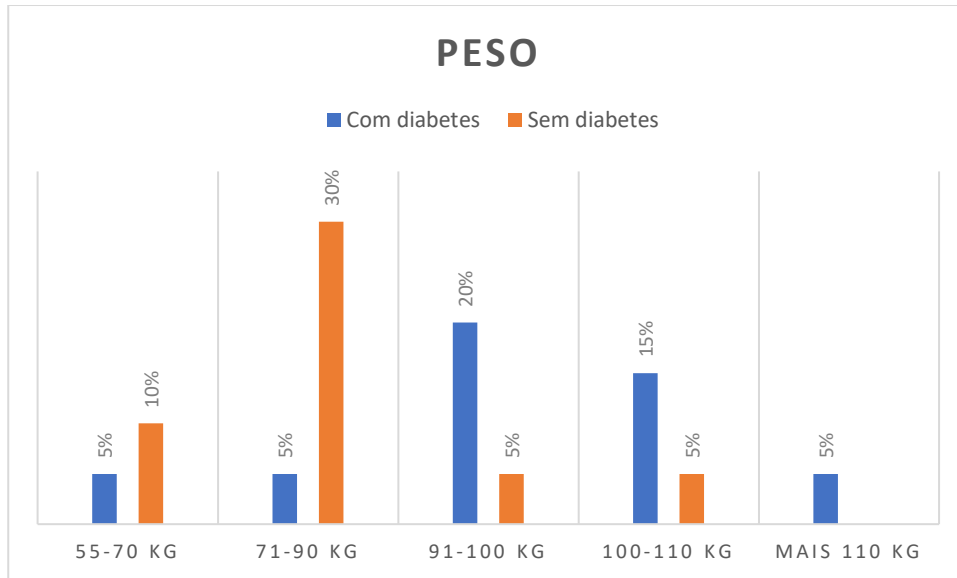
Tabela 05 – Avaliação Física dos sinais vitais dos voluntários não diabéticos

VOLUNTÁRIOS (Sem Diabetes)	PA	TEMPERATURA	GLICEMIA
Vn 01	140X70 mmHg	36,7	97 mg/dL
Vn 02	131X80 mmHg	37,0	96 mg/dL
Vn 03	130X90 mmHg	36,7	93 mg/dL
Vn 04	120X90 mmHg	37,2	99 mg/dL
Vn 05	120X70 mmHg	35,8	99 mg/dL
Vn 06	121x70 mmHg	36,5	87 mg/dL
Vn 07	130x90 mmHg	36,0	91 mg/dL
Vn 08	130x100 mmHg	36,6	89 mg/dL
Vn 09	130x110 mmHg	36,5	100 mg/dL
Vn 10	120x80 mmHg	36,7	92 mg/dL

Fonte: Autoria Própria

De acordo com a Tabela 05, dos voluntários sem diabetes apenas 20% dos participantes encontram-se com a pressão arterial estável, enquanto 80% apresentam a pressão arterial acima. No que se refere à temperatura, 100% a mantém em estado normal. Mas o fator glicêmico apresenta-se dentro parâmetro do valor de referência que estabelece a normalidade (70 a 100mg/dL) em todos os participantes.

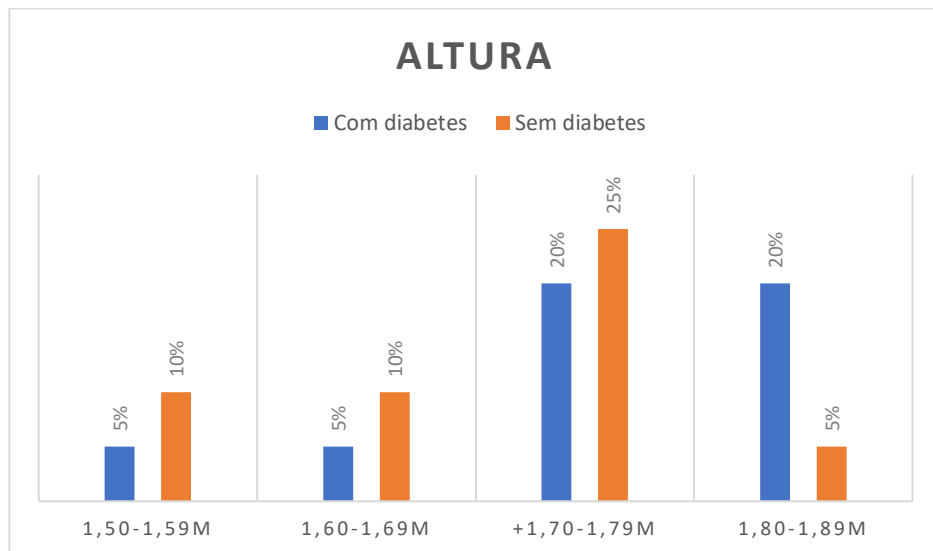
Figura 21 – Peso dos voluntários



Fonte: Autoria Própria

De acordo com a Figura 21, observam-se 5% dos voluntários com mais de 110 kg, 20% com peso entre 100 e 110kg, 25% com peso entre 91 e 110kg, 35% com peso entre 71 e 90 kg, e 15% com peso de 55 e 70kg.

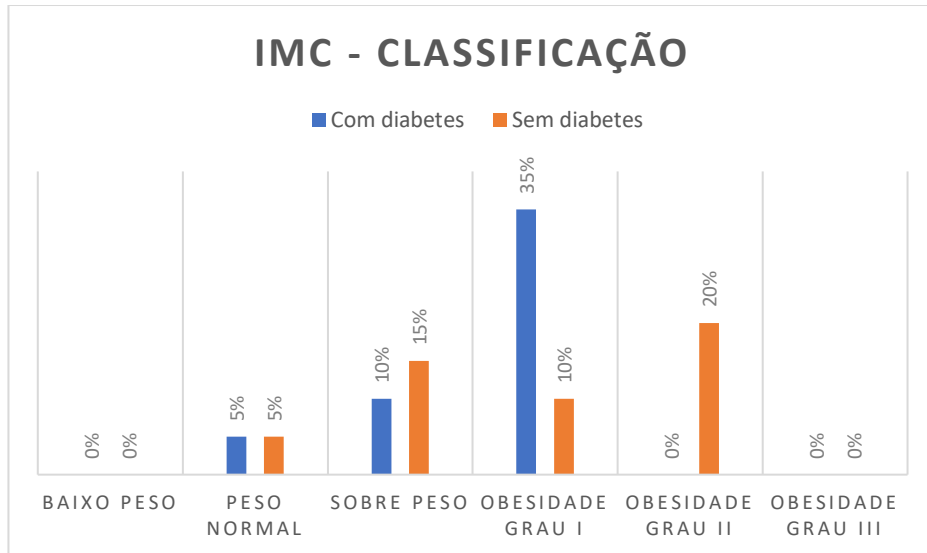
Figura 22 – Estatura dos voluntários



Fonte: Autoria Própria

De acordo com a medição da estatura dos voluntários exposto na Figura 22, pode-se afirmar que há 15% dos voluntários com altura entre 1,50-1,59m; 15% com altura entre 1,60-1,69m; 45% com altura entre 1,70-1,79m; 25% com altura entre 1,80-1,89m.

Figura 23 – Classificação do IMC de acordo com a OMS dos voluntários

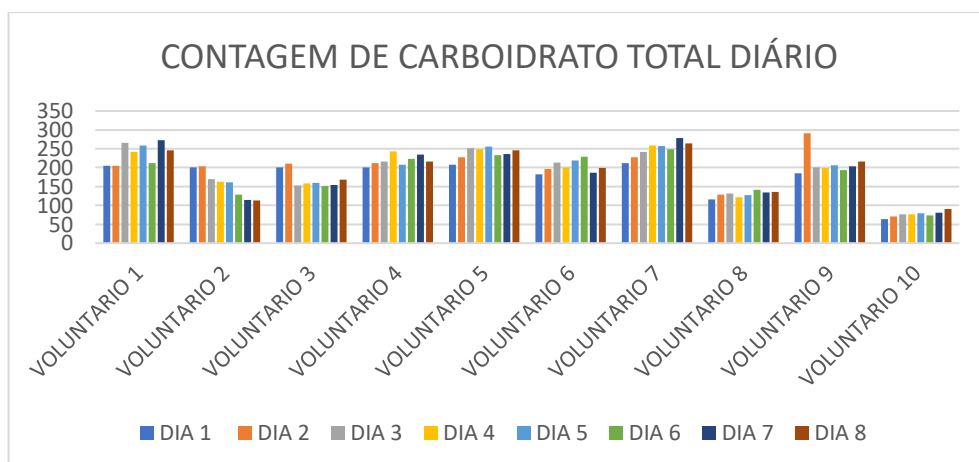


Fonte: Autoria própria

No Figura 23, dentre os voluntários com diabetes, têm-se 5% com peso normal, 10% com sobrepeso e 35% com obesidade grau 1. Dentre os voluntários sem diabetes, têm-se 5% com peso normal, 15% com sobrepeso, 10% com obesidade grau 1 e 20% com obesidade grau 2.

Compreendendo que o carboidrato é um dos nutrientes que mais afeta a glicemia e que 100% convertem-se em glicose (SANTOS et al, 2018), assim, fez-se necessário verificar a quantidade de carboidrato consumido diariamente pelos voluntários diabéticos e não diabéticos.

Figura 24 – Contagem de Carboidrato dos voluntários diabéticos

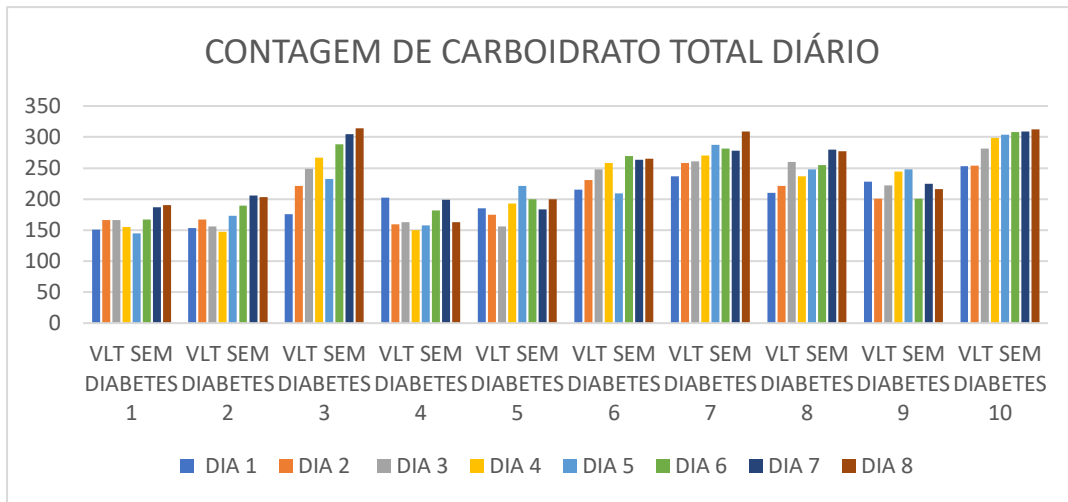


Fonte: Autoria própria

Dos voluntários diabéticos, observa-se no Figura 24 que o consumo varia entre 50 e 300g ou ml de carboidratos por dia, verificando uma frequência em 45%

dos participantes entre 200 e 250g ou ml de carboidratos. Esse padrão de consumo é a média diária consumida pelos brasileiros.

Figura 25 – Contagem de Carboidrato dos voluntários não diabéticos



Fonte: Autoria própria

Dos voluntários não diabéticos, observa-se no Figura 25 que o consumo varia entre 150 e 300 carboidratos ao dia, verificando uma frequência em 50% dos participantes entre 200 e 250 de carboidratos diários.

Pode-se observar que a contagem de carboidrato contou como um fator de controle glicêmico em apenas 5% do total de voluntários que participaram desta pesquisa. Este dado mostra a importância da predição dos dados glicêmicos neste estudo para um eventual controle desses níveis.

5.1 PREDIÇÃO DOS DADOS SISTÊMICOS

O modelo LSTM criado tem um total de 854,420 pesos treináveis. A tabela 06 apresenta a descrição do modelo.

Tabela 06 – Descrição do modelo LSTM com todos os voluntários

Camada (modelo)	Forma da Saída	Parâmetros #
Istm (LSTM)	(48,48)	9600
Istm 1 (LSTM)	(48, 200)	199200
Istm 2 (LSTM)	(48, 200)	320800
Istm 3 (LSTM)	(200)	320800
Dense (Dense)	(20)	4020
Total de Parâmetros: 854,420		
Parâmetros de treinamento: 854,420		
Sem Parâmetros de treinamento: 0		

O formato de saída representa a quantidade de unidades de entrada e unidades de saída. Na primeira camada, tem-se 48 entradas e 48 unidades de saída. A segunda camada contém 200 unidades de saída, e cada unidade está conectada às 48 unidades de saída. A terceira camada contém 200 unidades de saída e cada unidade está conectada às 48 unidades de saída. A quarta camada contém 200 unidades de saída e cada unidade está conectada às 200 unidades. A quinta camada contém 20 unidades, responsáveis pelas 20 observações preditas, e cada uma dessas unidades está conectada às 200 unidades de saída da quarta camada.

O resultado da validação cruzada usando o método *k-Fold* é apresentado na Tabela 07:

Tabela 07 – Resultado da validação cruzada dos dados, contendo os voluntários diabéticos e não diabéticos

Métrica	Valor médio	Desvio padrão	Taxa de acerto
MAE	0,20	0,01	80%
MSE	0,06	0,01	94%
RMSE	0,25	0,02	75%

Fonte: Autoria própria

É possível observar na Tabela 07 que os valores das métricas não variam muito entre a mudança das pastas, ficando entre 1% e 2%. Ademais, é possível observar que as métricas retornaram um bom resultado. É interessante ressaltar que a métrica MSE é sensível a ruídos por conta da elevação do erro ao quadrado. Um resultado de 94% significa que os valores preditos não se distanciaram muito dos valores reais.

Após a validação dos dados, fez-se o treinamento utilizando 70% dos dados, testou-se o modelo com a base de teste e obteve-se os seguintes resultados nas métricas definidas acima:

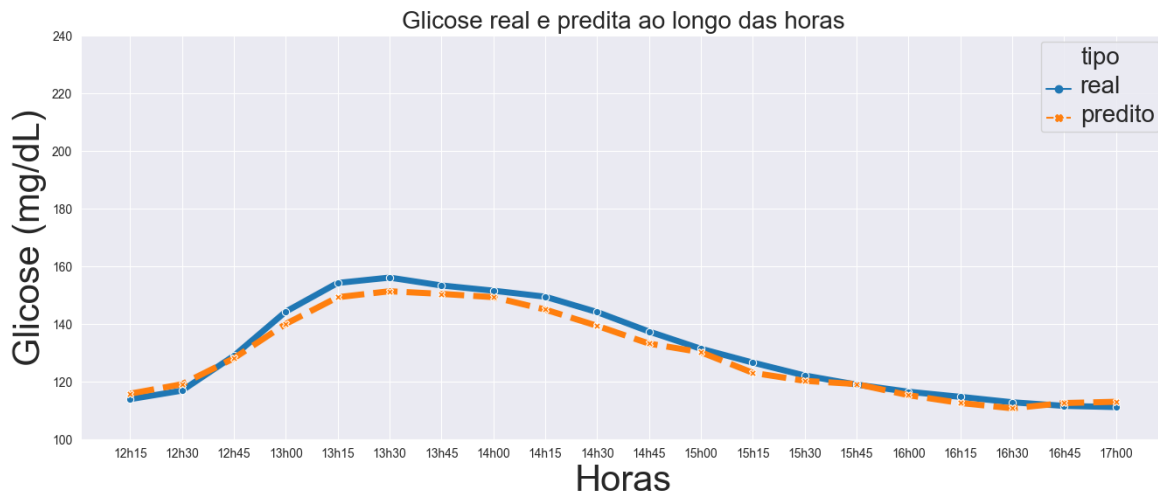
Tabela 08 – Métricas de Regressão

MÉTRICA	VALOR MÉDIO	TAXA DE ACERTO
MAE	0,195	80,462%
MSE	0,059	94,137%
RMSE	0,242	75,787%

Fonte: Autoria própria

Os resíduos e a variância do erro entre o valor real e predito se aproximam de 0. Esses resultados mostram que a média das predições se assemelha com a média dos valores reais, como é visto na Tabela 08, indicando uma boa performance média do modelo LSTM.

Figura 26 – Resultado do modelo LSTM treinado com todos os dados



Fonte: Autoria própria

É possível notar, no Figura 26, que o modelo conseguiu capturar o pico da glicose que acontece entre as 13:15 e 13:45 no grupo total. Esse pico coincide com os que foram observados no grupo de diabéticos e não diabéticos (Figura 11 e 12) e preditos pelo modelo. Observa-se também que a linha do modelo de predição acompanha o movimento glicêmico real, conseguindo antecipar 5 horas do possível nível glicêmico a partir de dados contínuos coletados de 12 horas antes.

Para fins de ilustração, delimitou-se esse trecho de 5 horas do predito com o intuito de demonstrar que o modelo ocorre de acordo com os dados glicêmicos coletados de todos os participantes.

Os dois modelos LSTM individuais, que foram criados para 1 voluntário diabético e 1 voluntário não diabético, têm um total de 864.840 pesos treináveis cada. A tabela 09 e 10 abaixo apresenta a descrição do modelo:

Tabela 09 – Descrição dos modelos LSTM treinados individualmente com os dados de um voluntário diabético

Camada (modelo)	Forma da Saída	Parâmetros #
Istm (LSTM)	(48,48)	9600
Istm_1 (LSTM)	(48, 200)	199200
Istm_2 (LSTM)	(200)	320800
Dense (Dense)	(20)	4020
Total de Parâmetros: 533.620		
Parâmetros de treinamento: 533.620		
Sem Parâmetros de treinamento: 0		

Fonte: Autoria própria

Tabela 10 – Descrição dos modelos LSTM treinados individualmente com os dados de um voluntário não diabético

Camada (modelo)	Forma da Saída	Parâmetros #
Istm (LSTM)	(48,48)	9600
Istm_1 (LSTM)	(48, 200)	199200
Istm_2 (LSTM)	(100)	120400
Dense (Dense)	(20)	2020
Total de Parâmetros: 331.220		
Parâmetros de treinamento: 331.220		
Sem Parâmetros de treinamento: 0		

Fonte: Autoria própria

Fez-se o treinamento utilizando 80% dos dados, testou-se o modelo com a base de teste e obtiveram-se os seguintes resultados nas métricas definidas nas tabelas 11 e 12:

Tabela 11 – Modelo dos voluntários diabéticos

MÉTRICAS	VALOR MÉDIO	TAXA DE ACERTO
MAE	0,102	89,755%
MSE	0,018	98,167%
RMSE	0,135	86,460%

Fonte: Autoria própria

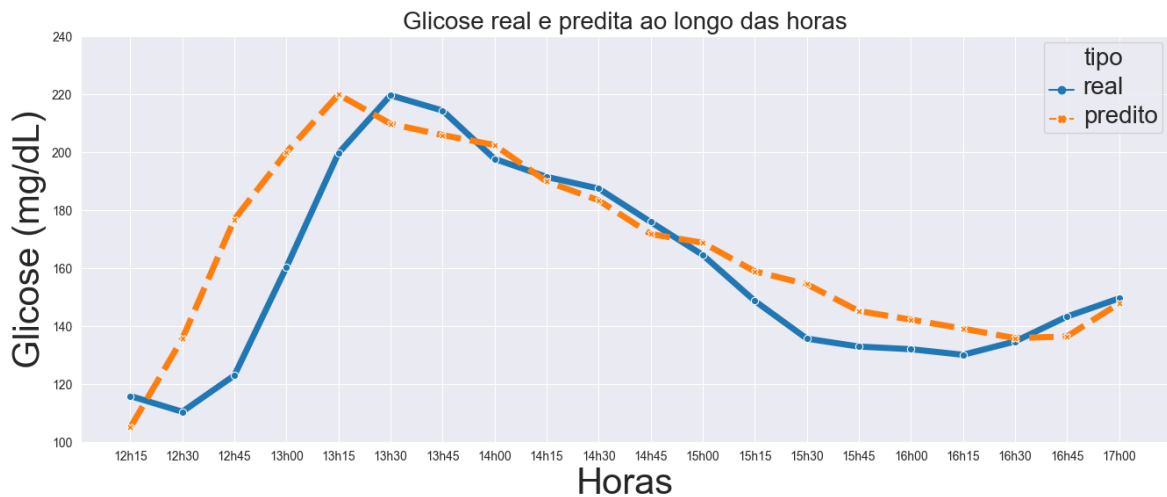
Tabela 12 – Modelo dos voluntários não diabéticos

MÉTRICAS	VALOR MÉDIO	TAXA DE ACERTO
MAE	0,233	76,720%
MSE	0,075	92,490%
RMSE	0,274	72,596%

Fonte: Autoria própria

É possível observar que a melhor predição teve o melhor desempenho para o modelo dos voluntários diabéticos, ou seja, ela é melhor que o modelo geral que leva todos em consideração e que o modelo personalizado dos não diabéticos.

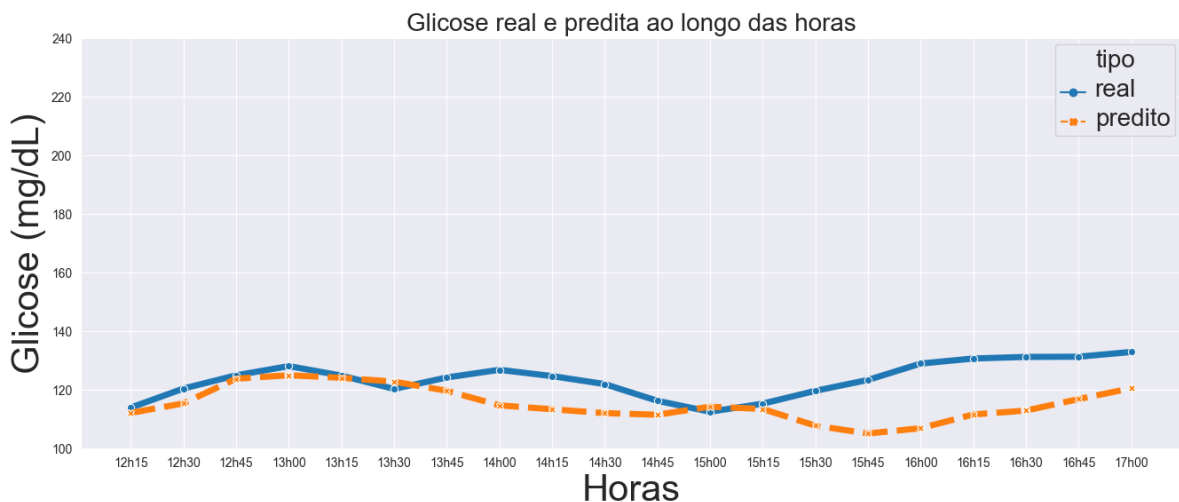
Figura 27 – Resultado do modelo LSTM treinado com os dados dos voluntários diabéticos



Fonte: Autoria própria

É possível notar no Figura 27 que a curva do valor predito de 220 mg/dL de glicose às 13:15 é o mesmo do valor real obtido às 13:30 pelo sensor. Os indicadores avaliativos mostram uma boa performance do modelo, bem como as tendências de subida e descida do valor da glicose.

Figura 28 – Resultado do modelo LSTM treinado com os dados dos voluntários não diabéticos



Fonte: Autoria própria

Os dados dos voluntários não diabéticos não possuem padrões que sejam tão previsíveis em relação aos diabéticos, ou seja, o modelo não foi capaz de captar todas as nuances necessárias para uma boa predição porque os dados não diabéticos não possuem variabilidade significativa para um modelo; diferente dos voluntários diabéticos, onde apresenta altos picos no valor da glicose.

6 DISCUSSÕES

A presente pesquisa trata de um modelo preditivo do nível glicêmico por monitoramento em tempo real em indivíduos portadores de diabetes mellitus tipo II, contando com 20 voluntários para o *corpus* do trabalho. Os dados coletados pelo sensor representaram 1.344 dados de glicemia observados por 1 voluntário, chegando a um total de 26.880 nos 20 voluntários que participaram deste estudo

Após essa coleta de dados por sensor CGM, fez-se um programa LSTM, utilizando o método k-fold, com cinco camadas para realizar uma predição da glicemia, fazendo uso de 70% dos dados para treino e 30% para teste.

Com base nos resultados apresentados pelo programa, foi possível criar um modelo para a predição da glicose ao longo do dia, captando as nuances necessárias para o objetivo deste trabalho. Os modelos apresentaram bons índices nas métricas escolhidas.

O modelo personalizado para um voluntário diabético performou melhor que os outros dois (modelo geral e o modelo individual não diabético). Dessa forma, pode-se notar que dados de um só indivíduo são o suficiente para criar um modelo de predição do valor da glicose. Assim, o modelo aprende as características particulares de cada um e aumenta a precisão na predição desse valor.

Vale ressaltar que o consumo de carboidrato diário entre 200 e 250g/ml pelos voluntários diabéticos e não diabéticos converge com os dados glicêmicos obtidos pelo programa. Pressupõe-se, assim, que os picos de glicemias observados tanto nos diabéticos quanto nos não diabéticos ocorreram, provavelmente, pelo excesso de consumo de carboidrato, principalmente ocorridas após as refeições diárias, no caso a pós-prandial.

É importante lembrar, de acordo com a Sociedade Brasileira de Diabetes (2016, p. 10), “o carboidrato, embora saudável, é o nutriente que tem maior efeito sobre a glicemia, já que a totalidade (100%) do que é ingerido transforma-se em glicose”; enquanto da proteína, um valor entre 30% e 60% pode ser transformado em glicose e da gordura, somente 10%. A contagem de carboidratos no Diabetes tipo 2, com controle apenas com dieta e atividade física ou com antidiabéticos orais, consiste em conhecer a quantidade de carboidratos a ser ingerida em cada refeição, mantendo esta quantidade sempre.

Os dados glicêmicos coletados performaram de forma positiva, pois os valores preditos acompanham o movimento glicêmico, sendo que o modelo preditivo do diabético foi o que melhor performou, convergindo com os dados obtidos da contagem de carboidratos, avaliação física e antropométrica, observados com os picos de glicemia, modos de vida dos voluntários e os carboidratos totais consumidos diariamente.

A maioria dos 20 voluntários possuem uma vida sedentária, não fazem atividade física diária nem uso de uma alimentação saudável e condizente com a doença em questão, estando a maioria acima do peso normal. O perfil dos voluntários encontra-se entre 30 e 65 anos, com renda familiar variando de 1 a mais de 6 salários, a maioria com ensino superior completo e trabalhando em empresa pública.

Essa observação coaduna com os estudos de Felipeti (2016) ao afirmar que o DM2, que deriva da combinação de deficiência relativa de insulina e resistência insulínica, normalmente está associada à obesidade, hipertensão e dislipidemia, acometendo principalmente as pessoas acima de 40 anos.

A incidência da DM2 ocorre mais comumente e de forma ascendente especialmente em países em desenvolvimento e relaciona-se com o hábito de vida ocidental. Felipetti (2016) explica que a industrialização ao favorecer um maior conforto para a população acaba levando às pessoas se exercitarem menos. As consequências são o aumento do peso, da gordura localizada predominantemente no abdome e maior resistência insulínica, caracterizando o diabético tipo 2.

De acordo com os dados obtidos dentre os diabéticos, torna-se comum haver hipertensão arterial e distúrbios visuais como observado nos questionários respondidos e confirmados também pela aferição da pressão arterial dos voluntários observados na avaliação antropométrica, além de estarem acima do peso normal.

Esses dados são ratificados por Malta (2019) em sua pesquisa quando aponta que a prevalência de DM2 é de 9,4% na população geral e se torna ainda mais significativa com o aumento da idade, cuja prevalência é de 22,6% na população maior de 60 anos. No que se refere à hipertensão arterial, tem-se uma prevalência de cerca de 24,0%, alcançando 60,9% na população idosa. (BRASIL, 2019).

Pode-se dizer, assim, que esse modo de vida converge com os dados glicêmicos captados pelo sensor de monitoramento de glicemia contínua. E, de modo geral, ser possível fazer a predição da glicose ao longo do dia e criar um modelo personalizado para cada indivíduo. Portanto, o modelo aprenderia com os dados de cada um e faria predições com uma precisão maior, atualizando sempre de acordo com novos dados.

7 CONCLUSÃO

Ao voltar ao problema inicial: *Qual a melhor forma de se obter uma predição de nível glicêmico dos sujeitos que possuem diabetes mellitus tipo II?*, observa-se ser possível obter um modelo preditivo do nível glicêmico de portadores de diabetes mellitus tipo II, utilizando modelos de inteligência artificial.

Como as setas de tendência do aparelho de monitoramento contínuo de glicemia apontam para um movimento glicêmico, tornou-se necessário não apenas o direcionamento, mas também uma possível predição glicêmica.

A diabetes mellitus configura-se no âmbito mundial como uma das principais doenças crônicas que atinge grande parte do contingente populacional. Ao notar o avanço da medicina associada à Inteligência Artificial, este projeto teve o objetivo de desenvolver um método de predição do nível glicêmico de portadores de diabetes mellitus tipo II utilizando técnicas de Inteligência Artificial (IA) a partir de séries temporais obtidos por sistema de monitoramento contínuo de glicose.

Para desenvolver o estudo em questão, foram aplicados formulários a fim de se obter características sociodemográficas e de hábitos de vida dos voluntário. Referente ao consumo diário dos carboidratos, fez-se o uso do aplicativo da contagem de carboidratos da sociedade brasileira de diabetes. Esses dados foram comparados aos dados captados e registrados pelo sensor do *freestyle libre* e descarregados via usb no computador.

Com base nos resultados apresentados, foi possível criar um modelo para a predição da glicose ao longo do dia, captando as nuances necessárias para o objetivo deste trabalho. Os modelos apresentaram bons índices nas métricas escolhidas.

O modelo personalizado para um voluntário diabético desempenhou melhor que os outros dois (modelo geral e o modelo individual não diabético). Dessa forma pode-se notar que dados de um só indivíduo são o suficiente para criar um modelo de predição do valor da glicose.

Com os dados obtidos por sensor CGM criou-se um programa LSTM, utilizando o método K-fold, com cinco camadas para realizar uma predição da glicemia, fazendo uso de 70% dos dados para treino e 30% para teste no modelo geral, e 80% dos dados para treino e 20% para teste no modelo individual. Com os

dados glicêmicos coletados dos 20 voluntários, obtiveram-se valores preditos que acompanharam o movimento glicêmico e convergindo com os dados obtidos da contagem de carboidratos, avaliação física e antropométrica, observados com os picos de glicemia, modos de vida dos voluntários e os carboidratos totais consumidos diariamente.

Assim, o modelo aprende as características particulares de cada um e aumenta a precisão na predição desse valor. De modo geral, é possível fazer a predição da glicose ao longo do dia, sendo possível criar um modelo personalizado para cada indivíduo. Portanto, dessa forma, o modelo aprenderia com os dados daquele indivíduo e faria predições com uma precisão maior, atualizando sempre de acordo com a geração de novos dados.

Apesar de alcançar o objetivo aqui proposto, a pesquisa contou com algumas limitações. O fato de estar em plena pandemia do COVID-19 proporcionou que a coleta de dados dos voluntários fosse aplicada com as devidas precauções que o momento solicitava, precisando aplicar o sensor e coletar os dados um a cada vez, para que não gerasse contaminação entre os indivíduos envolvidos. Outro fator foi devido os sensores serem fomentados pelo próprio pesquisador, o que levou a aplicação de apenas 20 voluntários; entretanto, suficientes para a coleta de dados que os sensores obtiveram para o programa aqui desenvolvido, haja vista possuir um número considerável de dados para a programação.

A pesquisa aqui em questão torna-se extremamente importante para diversas áreas da saúde, já que o indivíduo com diabetes tipo II encontra-se com vários agentes da saúde. Diante disso, o programa aqui em questão pode ajudar esses agentes como melhor conduzir medicações, uma dieta nutricional, atividades físicas, de acordo com a predição de níveis glicêmicos e, assim, evitando situações prejudiciais ao portador de diabetes mellitus tipo II. E também no âmbito da inteligência artificial, de séries temporais, já que o monitoramento de glicose contínuo capta os dados glicêmicos do indivíduo de 15 em 15 min, sendo assim fonte de pesquisa para desenvolvimento de aplicativo.

REFERÊNCIAS

ABADI, M. et al. **Tensorflow**: Large-scale machine learning on heterogeneous distributed systems. 2016.

ABBOTT. **Freestyle-Libre-3-Sensor-And-Pennies**. 28 de setembro de 2020. Disponível em: https://abbott.mediaroom.com/2020-09-28-Abbotts-FreeStyle-R-Libre-3-System-Receives-CE-Mark-Features-Worlds-Smallest-Thinnest-Sensor-with-Best-in-Class-Performance-at-the-Same-Low-Cost-for-People-with-Diabetes#assets_2429_124427-111:19442 Acesso em: 16 dez 2022.

ALBUQUERQUE, Reginaldo. **O fim das picadas**: novos recursos para avaliação dos níveis de glicose. Disponível em: <https://www.anad.org.br/o-fim-das-picadas-novos-recursos-para-a-avaliacao-dos-niveis-de-glicose/>. Acesso em: 02.jun 2022.

AMERICAN DIABETES ASSOCIATION. Diagnosis and Classification of Diabetes Mellitus. **Diabetes Care**, 29 (Suppl 1), p.43-48, 2006a.

AMERICAN DIABETES ASSOCIATION. Standards of Medical Care in Diabetes–2006. **Diabetes Care**, 29 (Suppl 1), p. 4-42, 2006b.

BAILEY, T et al. The Performance and Usability of a Factory-Calibrated Flash Glucose Monitoring System. **Diabetes Technology & Therapeutics**, n. 17, v. 11, p. 787-794, 2015.

BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **Journal of Machine Learning Research**, v. 13, n. Feb, p. 281–305, 2012.

BERRY, S.E.; VALDES, A.M.; DREW, D.A. et al. Human postprandial responses to food and potential for precision nutrition. **Nat Med**, n. 26, p. 964–973, 2020. Disponível em: <https://doi-org.ez14.periodicos.capes.gov.br/10.1038/s41591-020-0934-0>. Acesso em: 23 mar. 2022.

BERTONHI, Laura Gonçalves; DIAS, Juliana Chioda Ribeiro. Diabetes mellitus tipo 2: aspectos clínicos, tratamento e conduta dietoterápica. **Revista Ciências Nutricionais Online**, v.2, n.2, p.1-10, 2018. Disponível em: <https://unifafibe.com.br/revistasonline/arquivos/cienciasnutricionaisonline/sumario/62/18042018212025.pdf> Acesso: 16 dez 2022.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. **Diabetes Mellitus**. Brasília, 2006. (Cadernos de Atenção Básica, n. 16).

BRASIL. Ministério da Saúde. Secretaria de Atenção Básica à Saúde. Departamento de Atenção Básica. **Diabetes mellitus**. Brasília, DF: Ministério da Saúde, 2006. (Cadernos de Atenção Básica, n. 16)

BRASIL. Ministério da Saúde. Secretária de Políticas de Saúde. Departamento de Ações Programáticas Estratégicas. **Plano de reorganização da atenção à hipertensão arterial e ao diabetes mellitus**: hipertensão arterial e diabetes mellitus. Brasília: Ministério da Saúde, 2001.

BRASIL. Ministério da Saúde. **Vigitel Brasil 2018**: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2018. Brasília: Ministério da Saúde, 2019.

CARDOSO, H et al. Consenso Nacional para a Utilização do Sistema de Monitorização Flash da Glicose. **Revista Portuguesa de Diabetes**. n. 13, v. 4, p. 143-153, 2018.

CARVALHO, A. C. P. L. F., et al. **Inteligência Artificial—uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011. p. 45.

CHAWLA, M. et al. Consensus and recommendations on continuous glucose monitoring. **J. Diabetol.**, n.10, p. 4-14, 2019.

CHICO, A. et al. Clinical Approach to Flash Glucose Monitoring: An Expert Recommendation. **J Diabetes Sci Technol**, 2019.

CHEN, Zhenyu et al. Condition Monitoring of Wing Turbine Gearbox Bearing Based on Deep Learning Model. (Monitoramento da condição do rolamento da Caixa de engrenagens da turbina eólica com base no modelo de aprendizado profundo). Abril, 2019. **IEEE ACCESS**, v. 4, 2016.

DAKE, Mysid. **A simplified view of an artificial neural network**. 28 November 2006. Disponível em: https://pt.wikipedia.org/wiki/Ficheiro:Neural_network.svg Acesso em: 16 dez. 2022.

ELLAHHAM, Samer. Artificial Intelligence: The Future for Diabetes Care. **The American Journal of Medicine**, n. 133, p. 895–900, 2020.

EXTENSIO. **R. Eletr. de Extensão**, Florianópolis, v. 17, n. 37, p. 61 – 72, 2020.

FEDOTENKOVA, Mariia. **Extraction de composants multivariés des signaux cérébraux obtenus pendant l'anesthésie générale**. These (Doctorat) – l'Université de Lorraine, 2016.

FELIPETTI, Francielly Andressa et al. Prevalência de hipertensos e diabéticos cadastrados e acompanhados pelas unidades de saúde do município de Cascavel – Paraná. **Rev. APS.**, n. 19, v.1, p. 77 - 84 jan/ma. 2016.

FOKKERT, MJ et al. Performance of the Freestyle Libre Flash Glucose Monitoring System in Patients with Type 1 And 2 Diabetes Mellitus. **BMJ Open Diabetes Research and Care**, n. 5, 2017.

FREESTYLE Libre. Sistema Flash de Monitorización de Glucosa. **Manual del usuario**, 2014.

GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2008.

GINSBERG, BH. Factors affecting blood glucose monitoring: sources of errors in measurement. **J Diabetes SciTechnol**, v. 3, n. 4, p. 903-913, 2009.

HAAK, T et al. Flash Glucose-Sensing Technology as a Replacement for Blood Glucose Monitoring for the Management of Insulin-Treated Type 2 Diabetes: a Multicenter, Open-Label Randomized Controlled Trial. **Diabetes Ther**, n. 8, v. 1, p. 55-73, 2017.

HAY, L.C.; WILMSHURST, E.G., FULCHER, G. Unrecognized hypo- and hyperglycemia in well-controlled patients with type 2 diabetes mellitus: the results of continuous glucose monitoring. **Diabetes Technol Ther**, v. 5, n. 1, p. 19-26, 2003.

HEINEMAN, L; FRECKMANN, G. CGM versus FGM: or Continuous Glucose Monitoring is not Flash Glucose Monitoring. **Journal of Diabetes Science and Technology**, v.9, n.5, p. 947-950, 2009

HEINEMANN, L.; FRECKMANN, G. CGM versus FGM; or continuous glucose monitoring is not flash glucose monitoring. **J Diabetes Sci Technol.**, n. 9, v. 5, p. 947-50, 2015.

HEINEMANN, L; FRECKMANN, G. CGM Versus FGM; or Continuous Glucose Monitoring is not Flash Glucose Monitoring. **Journal of Diabetes Science and Technology**, n. 9, v. 5, p. 47–950, 2015.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural computation**, v. 9, n. 8, p. 1735-1780, 1997.

HOSS U., BUDIMAN E.S. Factory-Calibrated Continuous Glucose Sensors: The Science Behind the Technology. **Diabetes Technology & Therapeutics**, n. 19, v. 2, p. 44-50, 2017.

HOSS, U. et al. Feasibility of Factory Calibration for Subcutaneous Glucose Sensors in Subjects with Diabetes. **Journal of Diabetes Science and Technology**, n. 8, v. 1, p. 89–94, 2014.

IDF. **IDF Diabetes Atlas**. 8. ed. International Diabetes Federation. Brussels: IDF; 2017. 150 p.

JORDAN, M. I.; MITCHELL, T. M. Machine learning: Trends, perspectives, and prospects. **Science, American Association for the Advancement of Science**, v. 349, n. 6245, p. 255–260, 2015.

KING, P.; PEACOCK, I.; DONNELLY, R. The UK Prospective Diabetes Study (UKPDS): clinical and therapeutic implications for type 2 diabetes. **Br J Clin Pharmacol**, n. 48, v. 5, p. 643–8, 2001.

KLUYVER, T. et al. Jupyter Notebooks – a publishing format for reproducible computational workflows. In: LOIZIDES, F.; SCHMIDT, B. eds. **Positioning and Power in Academic Publishing: Players, Agents and Agendas**. 2016. p. 87–90

OHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **International joint Conference on artificial intelligence**. [S.l.: s.n.], 1995. v. 14, p. 1137–1145.

LOCA, Antonio Luiz da Silva. **Uma metodologia experimental para avaliar abordagens de aprendizado de máquina para diagnóstico de falhas com base em sinais de vibração**. Vitória, 126f. Dissertação (Mestrado em Informática) - Universidade Federal do Espírito Santo, Centro Tecnológico, 2020.

LYRA, Ruy; CAVALCANTI, Ney; SANTOS, Raul Dias (Org.). **Diabetes Mellitus: uma abordagem cardiovascular**. São Paulo: Editora Clannad, 2019.

LYRA, R; CARVALHO, N; COELHO, D. Definição, diagnóstico e classificação dos distúrbios no metabolismo dos hidratos de carbono. In: LYRA, Ruy; CAVALCANTI, Ney; SANTOS, Raul Dias (Org.). **Diabetes Mellitus: uma abordagem cardiovascular**. São Paulo: Editora Clannad, 2019.

MALTA, DC et al. Prevalência de diabetes mellitus determinada pela hemoglobina glicada na população adulta brasileira, Pesquisa Nacional de Saúde. **Rev Bras Epidemiol** [Internet], n. 22, v. 27, p. 1-13, 2019 [cited 2020 Jun 21]. Disponível em: https://www.scielo.br/scielo.php?pid=S1415-790X2019000300408&script=sci_arttext&lng=pt. Acesso em:

MATHUR, P. et al. (2020). Artificial Intelligence, Machine Learning, and Cardiovascular Disease. **Clinical Medicine Insights: Cardiology**. Disponível em: <https://doi.org/10.1177/1179546820927404>. Acesso em: 20 mar. 2022

MENENDEZ, Torres E. Monitorización de la glucemia en la diabetes. Perspectiva histórica y evolución tecnológica. **Av Diabetol**, v. 26 (Supl 1), p. 1-4, 2010.

MINAYO, Maria Cecilia de Souza. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 11. ed. São Paulo: Hucitec, 2008.

MIRA, G.S.; CANDIDO, L.M.B; YALE, J.F. Performance of glucometer used for self-monitoring blood glycaemia in type 1 diabetic patients. **Arq Bras Endocrinol Metabol**, n. 50, v. 3, p. 541-9, 2006.

PEDREGOSA et al. Scikit-learn: Machine Learning in Python. **JMLR**, n. 12, p. 2825-2830, 2011.

PYTHON SOFTWARE FOUNDATION. **Python Language Site**: Documentation, 2022. Página de documentação. Disponível em: <https://www.python.org/doc/>. Acesso em: 13 de mar. de 2022.

RIBAS, C.R.P. et al. Incidentes críticos no processo de ensino-aprendizagem em diabetes na perspectiva da equipe multiprofissional de saúde. **Rev Eletr Enferm**, n.10, v. 3, p. 747–55, 2008.

ROSA, M.Q.M. et al. Disease and economic burden of hospitalizations attributable to diabetes mellitus and its complications: a nationwide study in Brazil. **Int J Env Res Public Heal**, n.15, v. 2, p. 294, 2018.

ROSSANEIS, Mariana Angela et al. Fatores associados ao controle glicêmico de pessoas com diabetes mellitus. **Ciência & Saúde Coletiva** [online]. 2019, v. 24, n. 3, p. 997-1005. Disponível em: <<https://doi.org/10.1590/1413-81232018243.02022017>>. Acesso em: 16 dez. 2022.

RUMELHART, D., HINTON, G., WILLIAMS, R. Learning representations by back-propagating errors. **Nature**, n. 323, p. 533–536, 1986. Disponível em: <https://doi.org/10.1038/323533a0>. Acesso em: 20 de mar. de 2022.

SBD. **Manual de Contagem de Carboidratos para Pessoas com Diabetes**. São Paulo: Departamento de Nutrição da Sociedade Brasileira de Diabetes, 2016.

SBD. **Diabetes mellitus: uma abordagem cardiovascular 2019-2020**. São Paulo: Clannad Editora Científica, 2019.

SCUDILIO, Juliana. Parte III: Como utilizar modelos de machine Learning para reduzir o Churn. **Flai**. Data Science. 13 set. 2020. Disponível em: flai.com.br. Acesso em: 17 de dez. 2022.

SILVA, I.N.; SPATTI D.; FLAUZINO, R. **Redes neurais artificiais para engenharia e ciências aplicadas**: curso prático. São Paulo: Artliber Editora Ltda, 2010.

STEFENON, Paula. **Métodos para avaliação e acompanhamento de pacientes candidatos a terapia de substituição de células-beta**. Porto Alegre, 67f. Dissertação (Mestrado em Ciências Médicas: Endocrinologia) – Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, 2019.

TANVEER, M. et al. Machine Learning Techniques for the Diagnosis of Alzheimer's Disease: A Review. **ACM Trans. Multimedia Comput. Commun. Appl.** 16, 1s, Article 30, 35 pages, Jan. 2020. Disponível em: <https://doi.org/10.1145/3344998>. Acesso em: 20 mar. 2022

THE DIABETES Control and Complications. Research Group, The Diabetes Control and Complications Trial Research Group. Effect of intensive therapy on the development and progression of diabetic nephropathy in the Diabetes Control and Complications Trial. **Kidney Int.** n.47, v. 6, p.1703–20, 1995.

VIEIRA, Suzana. **Monitorização contínua da glicose – sistema Flash: tempo no alvo e outras informações.** 2019. Disponível em: <https://drasuzanavieira.med.br/2017/06/16/monitorizacao-continua-da-glicose-tempo-no-alvo/> Acesso em: 21 dez 2022.

WANG, T.; LI, W. (V.) Previsão de glicose no sangue usando variantes LSTM no contexto do sistema de pâncreas artificial de código aberto. In: 53^a CONFERÊNCIA INTERNACIONAL DO HAVAÍ SOBRE CIÊNCIAS DO SISTEMA, 7 a 10 de janeiro de 2020, Maui, Havaí. Disponível em <http://centaur.reading.ac.uk/86409/>. Acesso em: 16 dez. 2022.

APÊNDICES

APÊNDICE A – CARACTERIZAÇÃO DOS PARTICIPANTES DA PESQUISA

Sexo: () Feminino () Masculino

Idade: () 30 a 40 anos () 41 a 50 anos () 51 a 60 anos
() 61 a 60 anos () Mais de 60 anos

Situação conjugal: () casado () solteiro () união estável () divorciado ()
viúvo

Renda familiar (salários mínimos): () 1 a 3 () 4 a 6 () mais de 6

Nível de formação: () Analfabeto () Fundamental () Ensino Médio ()
Superior () Pós graduação

Trabalha em empresa pública? () Sim, qual área _____
() Não

Trabalha em empresa privada? () Sim, qual área _____
() Não

Trabalha como autônomo? () Sim, qual área _____
() Não

Tempo de trabalho: () Menos de 1 ano () 1 a 5 anos () 6 a 10 anos ()

Carga Horária: () 20 horas () 40 horas

Turno em que trabalha: () Manhã () Tarde () Noite .

APÊNDICE B – Mapeamento da situação de saúde do participante

Prática de atividade física:

sim Qual? _____ não

Em seu tempo livre, você possui algum lazer:

sim Qual? _____ não

Ingere álcool:

sim não

Se sim, qual a periodicidade?

1 vez na semana 2 vezes na semana 3 vezes na semana

4 vez na semana mais de 4 vezes na semana

Tabagista:

sim. Número de maços por dia _____ não

Após 30 anos teve alguma patologia específica?

sim não

Se positivo, qual?

hipertensão

diabetes

problemas cardiovasculares

disfunções musculoesqueléticas

outros... Qual? _____

Distúrbios visuais:

sim não

Distúrbios auditivos:

sim Quais _____ não

Uso de medicação:

sim não

Já fez alguma cirurgia?

sim Motivo(s) _____ não

APENDICE C – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Eu, _____, estou sendo convidado a participar do estudo **Modelo Preditivo do nível glicêmico por monitoramento flash em indivíduos portadores de diabetes mellitus**, com objetivo de desenvolver um método utilizando técnicas de Inteligência Artificial a partir de series temporais dos níveis glicêmicos obtidos por tecnologia flash. A minha participação no referido estudo será no sentido de *responder* perguntas relacionadas à sua saúde outro sobre a caracterização dos participantes e o preenchimento do aplicativo de contagem de carboidratos.

Fui alertado de que, da pesquisa a se realizar, posso esperar alguns benefícios, tais como: Resultado da glicemia com sensor, explicação do questionário sobre sua saúde e treinamento do uso do aplicativo de contagem de carboidratos. Recebi, por outro lado, os esclarecimentos sobre os possíveis desconfortos e riscos decorrentes do estudo como: A realização deste estudo não irá expor você a qualquer tipo de risco ou constrangimento. Entretanto, você poderá sentir algum desconforto relacionado ao tema de pesquisa ao responder o questionário. Poderá desistir de participar a qualquer momento. Além disso, se você sentir necessidade, poderá ser encaminhado ao Serviço Social e Médico da instituição em estudo, sem qualquer custo na consulta. Estou ciente de que minha privacidade será respeitada, ou seja, meu nome ou qualquer outro dado ou elemento que possa, de qualquer forma, me identificar, será mantido em sigilo. Também fui informado de que posso me recusar a participar do estudo ou retirar meu consentimento a qualquer momento, sem precisar justificar e por desejar sair da pesquisa, não sofrerei qualquer prejuízo.

Os pesquisadores envolvidos com o referido projeto do estudo estão vinculados ao Programa de Pós-Graduação em Engenharia Biomédica, da Universidade Brasil Campus Itaquera SP, desenvolvido pelo doutorando Marcelo Henrique de Vasconcelos Mourão, sob a orientação do Prof. Dr. Marcello Magri Amaral, e com eles poderei manter contato pelos telefones (98) 98135-6312 e (11) 99780-1966). É assegurada a assistência durante toda pesquisa, bem como me é garantido o livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo, tendo sido orientado quanto ao teor de todo o aqui mencionado e compreendido a natureza e o objetivo do estudo, manifesto meu livre consentimento em participar, estando totalmente ciente de que não há nenhum valor econômico, a receber ou a pagar, por minha participação.

Em caso de reclamação ou qualquer tipo de denúncia sobre este estudo devo ligar para o Comitê de Ética em Pesquisa da Universidade Brasil (11) 20700025 sediado a Rua Carolina Fonseca, 235, Jd Santana, SP- Capital ou mandar um *e-mail* para comite.etica.sp@universidadebrasil.edu.br

São Luís – Ma, 20 de outubro de 2020.

Nome e assinatura do sujeito da pesquisa

Nome(s) e assinatura(s) do(s) pesquisador(es) responsável(veis)

APÊNDICE D – Modelo LSTM geral

Modelo LSTM Geral

October 9, 2022

```
[1]: import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'

import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import matplotlib.lines as mlines
import seaborn as sns
sns.set_style("darkgrid")

from statistics import mean

## Keras lib
import tensorflow as tf
from keras.models import Model, Sequential
from keras.utils import np_utils
from keras.layers import LSTM, Dense

## Scikit-learn
from sklearn.model_selection import train_test_split, KFold
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

tf.random.set_seed(42)
tf.compat.v1.logging.set_verbosity(tf.compat.v1.logging.ERROR)

N_STEPS_IN = 48
N_STEPS_OUT = 20
N_STEPS_MAX = N_STEPS_IN + N_STEPS_OUT
N_FEATURES = 1
```

1 Leitura dos dados

```
[2]: SERIES_PATH = "data/data_series"
```

```
[3]: def normalize(array):
      scaler = MinMaxScaler(feature_range=(0, 1))
      array_scaled = scaler.fit_transform(array.reshape(-1, 1))

      return array_scaled
```

```
[4]: file_names = os.listdir(SERIES_PATH)
      sequences = []

      scaler = MinMaxScaler(feature_range=(0, 1))

      for file_name in file_names:
          seq = np.genfromtxt(f'{SERIES_PATH}/{file_name}')
          # convert to [rows, columns] structure
          seq = seq.reshape((-1, 1))
          # normalize the sequence
          seq = scaler.fit_transform(seq)
          sequences.append(seq)

      print(len(sequences))
```

216

2 Divisão em dados de treinamento e de teste

```
[5]: X, y = [], []

      for i in range(len(sequences)):
          X.append(sequences[i][:N_STEPS_IN])
          y.append(sequences[i][N_STEPS_IN:N_STEPS_MAX])

      X, y = np.array(X), np.array(y)
```

```
[6]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.30,
      ↪random_state=42)
```

3 Criação do modelo LSTM

3.1 Parâmetros do modelo

```
[7]: epochs = 70
      activation = 'tanh'
      optimizer = 'adam'
      loss = 'mse'
```

3.2 Criação do modelo

```
[8]: tf.keras.backend.clear_session()

def create_model():
    model = Sequential()
    model.add(LSTM(N_STEPS_IN, activation=activation, return_sequences=True,
    ↪input_shape=(N_STEPS_IN, N_FEATURES)))
    model.add(LSTM(200, activation=activation, return_sequences=True,
    ↪input_shape=(200, 200)))
    model.add(LSTM(200, activation=activation, return_sequences=True))
    model.add(LSTM(200, activation=activation))
    model.add(Dense(N_STEPS_OUT))
    model.compile(optimizer=optimizer, loss=loss, metrics=["mae", "mse"])
    return model

model = create_model()
model.summary()
```

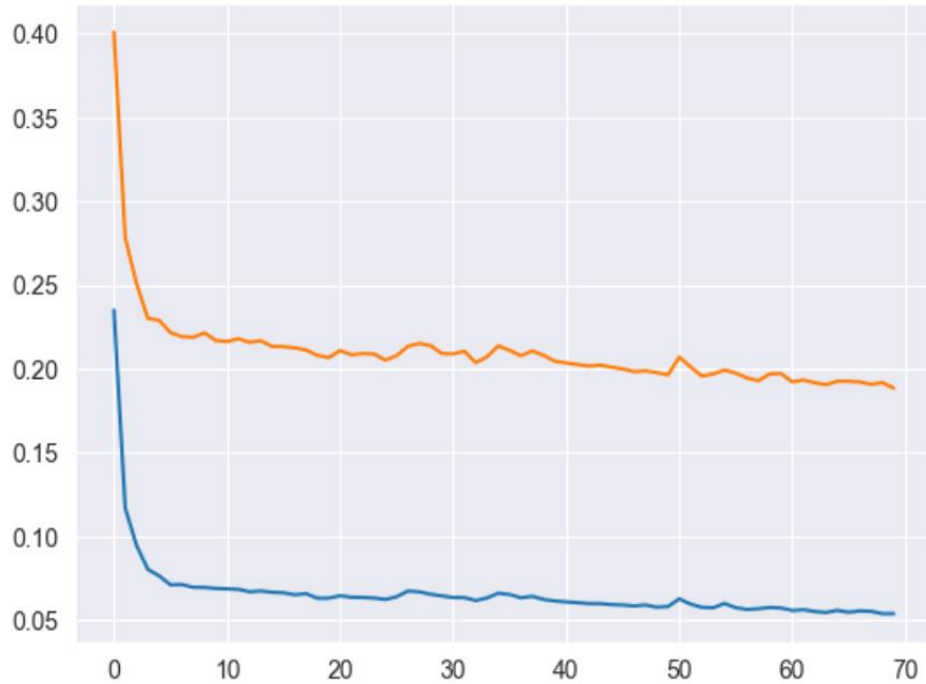
Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 48, 48)	9600
lstm_1 (LSTM)	(None, 48, 200)	199200
lstm_2 (LSTM)	(None, 48, 200)	320800
lstm_3 (LSTM)	(None, 200)	320800
dense (Dense)	(None, 20)	4020

=====
Total params: 854,420
Trainable params: 854,420
Non-trainable params: 0
=====

3.3 Treinamento do modelo

```
[9]: history = model.fit(X_train, y_train, epochs=epochs, verbose=0)
plt.plot(history.history['mse'])
plt.plot(history.history['mae'])
plt.show()
```



4 Teste do modelo criado

```
[10]: minutes = ["00", "15", "30", "45"]

def plot_result(y_test, y_pred):
    n = N_STEPS_OUT//4
    x_data = []
    n = N_STEPS_OUT//4
    x_data = ["12h15", "12h30", "12h45"]
    for i in range(n-1):
        for j in range(4):
            x = str((i+13)+"h"+str(minutes[j]))
            x_data.append(x)
    x_data.append("17h00")
    df = pd.DataFrame({"horas": x_data, "real": y_test, "predito": y_pred})

    data = pd.melt(df, id_vars=['horas'], value_vars=['real', "predito"],
                   var_name='tipo', value_name='glicose (mg/dL)')

    plt.figure(figsize=(16,6))
```

```

g = sns.lineplot(data=data, x="horas", y="glicose (mg/dL)", hue="tipo",
↳markers=True, style="tipo", linewidth = 5)
g.set_xlabel("Horas", fontsize = 30)
g.set_ylabel("Glicose (mg/dL)", fontsize =30)
g.set_ylim([100, 240])
plt.setp(g.get_legend().get_texts(), fontsize='20')
plt.setp(g.get_legend().get_title(), fontsize='20')
plt.title(label = "Glicose real e predita ao longo das horas", fontsize =
↳20)
plt.savefig('imgs/resultado_lstm.png')
plt.show()

def regression_report(y_true, y_pred):
    error = y_true - y_pred
    metrics = [
        ('mean absolute error', mean_absolute_error(y_true, y_pred)),
        ('mean squared error', mean_squared_error(y_true, y_pred)),
        ('root mean squared error', mean_squared_error(y_true, y_pred)**0.5)
    ]
    print('Metrics for regression:')
    for metric_name, metric_value in metrics:
        print(f'{metric_name:>25s}: {metric_value: >20.3f} {(1 -
↳metric_value)*100: >20.3f}%')

```

```

[11]: y_pred = model.predict(X_test)
regression_report(y_test.reshape(y_pred.shape[0], y_pred.shape[1]), y_pred)

```

3/3 [=====] - 1s 42ms/step

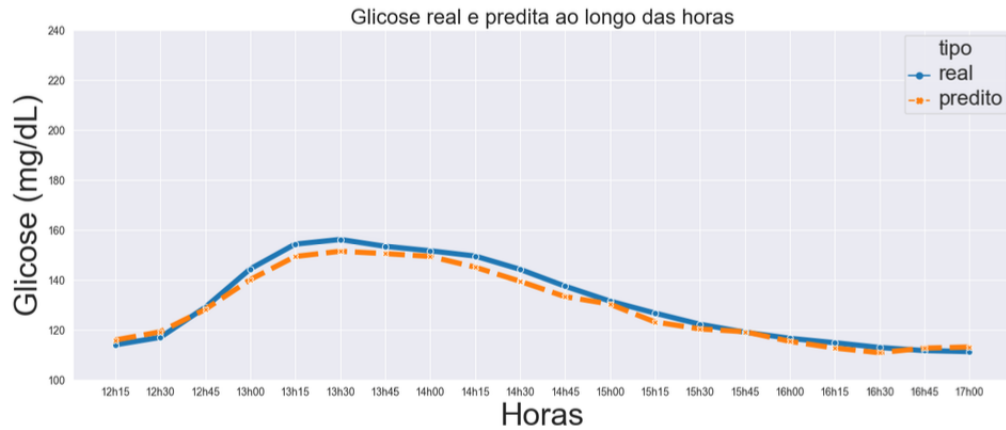
Metrics for regression:

mean absolute error:	0.195	80.462%
mean squared error:	0.059	94.137%
root mean squared error:	0.242	75.787%

```

[12]: y_test_ = scaler.inverse_transform(np.mean(y_test, axis=0))
y_pred_ = scaler.inverse_transform(np.mean(y_pred, axis=0).reshape(-1, 1))
plot_result(y_test_.reshape(1, -1)[0], y_pred_.reshape(1, -1)[0])

```



```
[13]: # Salvando o modelo
fname = "modelo_geral.h5"
model.save(fname)
```

5 Validação do modelo

```
[17]: performance = {
    "mae": {},
    "mse": {},
    "rmse": {}
}
```

```
[18]: k = 5
kfold = KFold(n_splits=k, shuffle=True)
print(X.shape, y.shape)
print(X.reshape(216,48).shape)
```

```
(216, 48, 1) (216, 20, 1)
(216, 48)
```

```
[ ]: fold = 1
for train, val in kfold.split(X, y):
    tf.keras.backend.clear_session()

    X_train = X[train]
    y_train = y[train]
    X_val = X[val]
    y_val = y[val]

    print(f"Pasta {fold} de {k}")
```

```

## Creating the model
model = create_model()

## Training the model
history = model.fit(X_train, y_train, epochs=epochs, verbose=1)
plt.plot(history.history['mse'])
plt.plot(history.history['mae'])
plt.show()

## Testing the model
result = model.evaluate(X_val, y_val)
performance["mse"][fold] = result[0]
performance["rmse"][fold] = result[0]**0.5
performance["mae"][fold] = result[1]
print(f'> Pasta {fold} - RMSE: {performance["rmse"][fold]}%')
fold += 1

```

```

[17]: output = {}

mean_mse = round(np.mean(list(performance["mse"].values())),2)
pct_mse = (1 - mean_mse) * 100
std_mse = round(np.std(list(performance["mse"].values())),2)
output["mse"] = str(mean_mse)
output["pct_mse"] = str(pct_mse) + "%"
output["std_mse"] = str(std_mse)

mean_rmse = round(np.mean(list(performance["rmse"].values())),2)
pct_rmse = (1 - mean_rmse) * 100
std_rmse = round(np.std(list(performance["rmse"].values())),2)
output["rmse"] = str(mean_rmse)
output["pct_rmse"] = str(pct_rmse) + "%"
output["std_rmse"] = str(std_rmse)

mean_mae = round(np.mean(list(performance["mae"].values())),2)
pct_mae = (1 - mean_mae) * 100
std_mae = round(np.std(list(performance["mae"].values())),2)
output["mae"] = str(mean_mae)
output["pct_mae"] = str(pct_mae) + "%"
output["std_mae"] = str(std_mae)

```

```
[18]: output
```

```

[18]: {'mse': '0.06',
      'pct_mse': '94.0%',
      'std_mse': '0.0',
      'rmse': '0.25',

      'pct_rmse': '75.0%',
      'std_rmse': '0.01',
      'mae': '0.2',
      'pct_mae': '80.0%',
      'std_mae': '0.01'}

```

APÊNDICE E- Modelo LSTM diabético

Modelo LSTM Diabético

October 9, 2022

```
[1]: import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'

import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import matplotlib.lines as mlines
import seaborn as sns
sns.set_style("darkgrid")

from statistics import mean

## Keras lib
import tensorflow as tf
from keras.models import Model, Sequential
from keras.utils import np_utils
from keras.layers import LSTM, Dense

## Scikit-learn
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

tf.random.set_seed(42)
tf.compat.v1.logging.set_verbosity(tf.compat.v1.logging.ERROR)

N_STEPS_IN = 48
N_STEPS_OUT = 20
N_STEPS_MAX = N_STEPS_IN + N_STEPS_OUT
N_FEATURES = 1
```

1 Leitura dos dados

```
[2]: SERIES_PATH = "data/diabetico"
```

```
[3]: def normalize(array):

    scaler = MinMaxScaler(feature_range=(0, 1))
    array_scaled = scaler.fit_transform(array.reshape(-1, 1))

    return array_scaled
```

```
[4]: file_names = os.listdir(SERIES_PATH)
sequences = []

scaler = MinMaxScaler(feature_range=(0, 1))

for file_name in file_names:
    if file_name[:2] == "20":
        seq = np.genfromtxt(f'{SERIES_PATH}/{file_name}')
        # convert to [rows, columns] structure
        seq = seq.reshape((-1, 1))
        # normalize the sequence
        seq = scaler.fit_transform(seq)
        sequences.append(seq)

print(len(sequences))
```

2 Divisão em dados de treinamento e de teste

```
[5]: X, y = [], []

for i in range(len(sequences)):
    X.append(sequences[i][:N_STEPS_IN])
    y.append(sequences[i][N_STEPS_IN:N_STEPS_MAX])

X, y = np.array(X), np.array(y)
```

```
[6]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20,
↳ random_state=42)
```

3 Criação do modelo LSTM

3.1 Parâmetros do modelo

```
[7]: epochs = 70
activation = 'tanh'
optimizer = 'adam'
loss = 'mse'
```

3.2 Criação do modelo

```
[8]: tf.keras.backend.clear_session()

model = Sequential()

model.add(LSTM(N_STEPS_IN, activation=activation, return_sequences=True,
↳input_shape=(N_STEPS_IN, N_FEATURES)))
model.add(LSTM(200, activation=activation, return_sequences=True))
model.add(LSTM(200, activation=activation))
model.add(Dense(N_STEPS_OUT))
model.compile(optimizer=optimizer, loss=loss, metrics=["mse", "mae"])
model.summary()
```

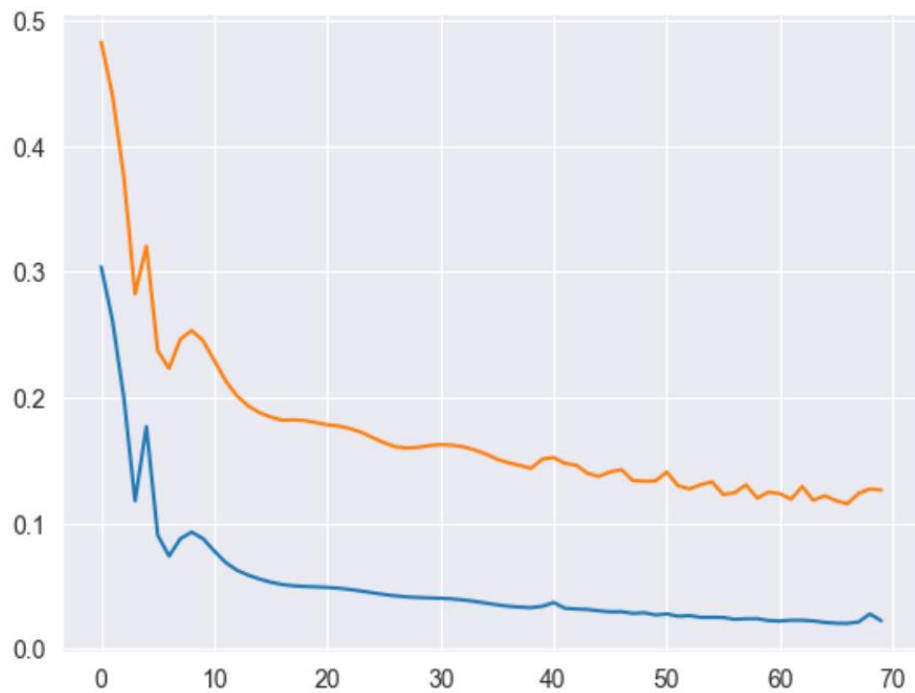
Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 48, 48)	9600
lstm_1 (LSTM)	(None, 48, 200)	199200
lstm_2 (LSTM)	(None, 200)	320800
dense (Dense)	(None, 20)	4020

=====
Total params: 533,620
Trainable params: 533,620
Non-trainable params: 0
=====

3.3 Treinamento do modelo

```
[9]: history = model.fit(X_train, y_train, epochs=epochs, verbose=0)
plt.plot(history.history['mse'])
plt.plot(history.history['mae'])
plt.show()
```



4 Teste do modelo criado

```
[10]: minutes = ["00", "15", "30", "45"]

def plot_result(y_test, y_pred):
    n = N_STEPS_OUT//4
    x_data = []
    n = N_STEPS_OUT//4
    x_data = ["12h15", "12h30", "12h45"]
    for i in range(n-1):
        for j in range(4):
            x = str((i+13))+ "h"+str(minutes[j])
            x_data.append(x)
    x_data.append("17h00")
    df = pd.DataFrame({"horas": x_data, "real": y_test, "predito": y_pred})

    data = pd.melt(df, id_vars=['horas'], value_vars=['real', "predito"],
        var_name='tipo', value_name='glicose (mg/dL)')

    plt.figure(figsize=(16,6))
```

```

g = sns.lineplot(data=data, x="horas", y="glicose (mg/dL)", hue="tipo",
↳markers=True, style="tipo", linewidth = 5)
g.set_xlabel("Horas", fontsize = 30)
g.set_ylabel("Glicose (mg/dL)", fontsize =30)
g.set_ylim([100, 240])
plt.setp(g.get_legend().get_texts(), fontsize='20')
plt.setp(g.get_legend().get_title(), fontsize='20')
plt.title(label = "Glicose real e predita ao longo das horas", fontsize =
↳20)
plt.savefig('imgs/resultado_lstm_individual_diabetico.png')

plt.show()

def regression_report(y_true, y_pred):

    error = y_true - y_pred

    metrics = [
        ('mean absolute error', mean_absolute_error(y_true, y_pred)),
        ('mean squared error', mean_squared_error(y_true, y_pred)),
        ('root mean squared error', mean_squared_error(y_true, y_pred)**0.5)
    ]

    print('Metrics for regression:')
    for metric_name, metric_value in metrics:
        print(f'{metric_name:>25s}: {metric_value: >20.3f} {(1 -
↳metric_value)*100: >20.3f}%')

```

```

[11]: y_pred = model.predict(X_test)
regression_report(y_test.reshape(y_pred.shape[0], y_pred.shape[1]), y_pred)

```

```
1/1 [=====] - 1s 1s/step
```

```
Metrics for regression:
```

mean absolute error:	0.102	89.756%
mean squared error:	0.018	98.167%
root mean squared error:	0.135	86.460%

```

[12]: y_test_ = scaler.inverse_transform(np.mean(y_test, axis=0))
y_pred_ = scaler.inverse_transform(np.mean(y_pred, axis=0).reshape(-1, 1))

plot_result(y_test_.reshape(1, -1)[0], y_pred_.reshape(1, -1)[0])

```

APÊNDICE F – Modelo LSTM Não diabético

Modelo LSTM Não Diabético

October 9, 2022

```
[1]: import os
os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'

import numpy as np
import pandas as pd

import matplotlib.pyplot as plt
import matplotlib.lines as mlines
import seaborn as sns
sns.set_style("darkgrid")

from statistics import mean

## Keras lib
import tensorflow as tf
from keras.models import Model, Sequential
from keras.utils import np_utils
from keras.layers import LSTM, Dense

## Scikit-learn
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

tf.random.set_seed(42)
tf.compat.v1.logging.set_verbosity(tf.compat.v1.logging.ERROR)

N_STEPS_IN = 48
N_STEPS_OUT = 20
N_STEPS_MAX = N_STEPS_IN + N_STEPS_OUT
N_FEATURES = 1
```

1 Leitura dos dados

```
[2]: SERIES_PATH = "data/nao_diabetico"
```

```
[3]: def normalize(array):

    scaler = MinMaxScaler(feature_range=(0, 1))
    array_scaled = scaler.fit_transform(array.reshape(-1, 1))

    return array_scaled
```

```
[4]: file_names = os.listdir(SERIES_PATH)
sequences = []

scaler = MinMaxScaler(feature_range=(0, 1))

cont = 0
for file_name in file_names:
    if cont == 8:
        break
    if file_name[:2] == "21":
        seq = np.genfromtxt(f'{SERIES_PATH}/{file_name}')
        # convert to [rows, columns] structure
        seq = seq.reshape((-1, 1))
        # normalize the sequence
        seq = scaler.fit_transform(seq)
        sequences.append(seq)
        cont += 1

print(len(sequences))
```

8

2 Divisão em dados de treinamento e de teste

```
[5]: X, y = [], []

for i in range(len(sequences)):
    X.append(sequences[i][:N_STEPS_IN])
    y.append(sequences[i][N_STEPS_IN:N_STEPS_MAX])

X, y = np.array(X), np.array(y)
```

```
[6]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20,
↪random_state=42)
```

3 Criação do modelo LSTM

3.1 Parâmetros do modelo

```
[7]: epochs = 70
      activation = 'tanh'
      optimizer = 'adam'
      loss = 'mse'
```

3.2 Criação do modelo

```
[8]: tf.keras.backend.clear_session()

      model = Sequential()

      model.add(LSTM(N_STEPS_IN, activation=activation, return_sequences=True,
                    →input_shape=(N_STEPS_IN, N_FEATURES)))
      model.add(LSTM(200, activation=activation, return_sequences=True))
      model.add(LSTM(200, activation=activation))
      model.add(Dense(N_STEPS_OUT))
      model.compile(optimizer=optimizer, loss=loss, metrics=["mse", "mae"])
      model.summary()
```

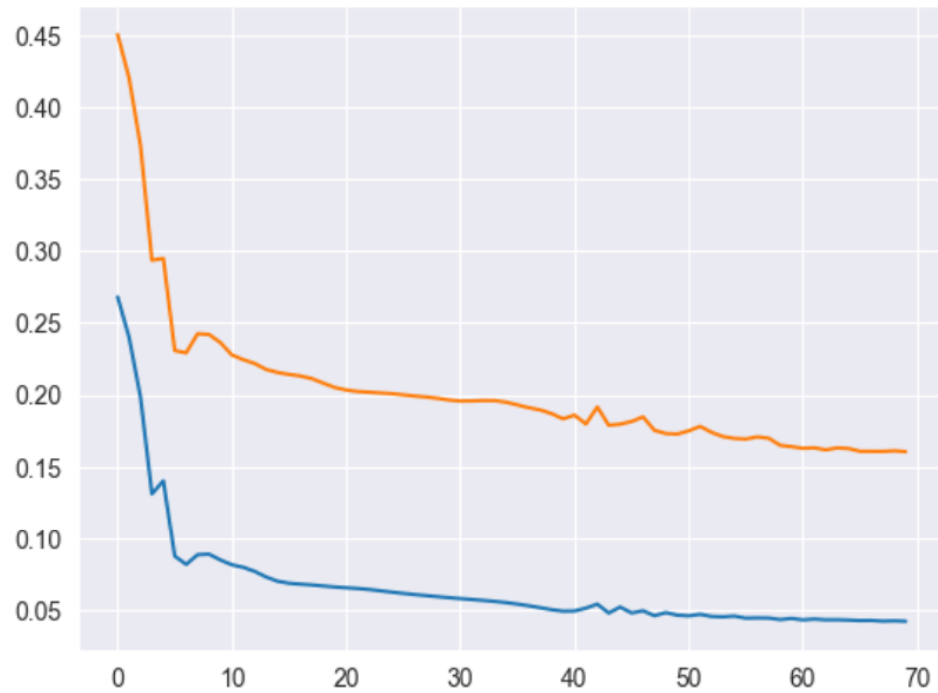
Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 48, 48)	9600
lstm_1 (LSTM)	(None, 48, 200)	199200
lstm_2 (LSTM)	(None, 200)	320800
dense (Dense)	(None, 20)	4020

=====
Total params: 533,620
Trainable params: 533,620
Non-trainable params: 0
=====

3.3 Treinamento do modelo

```
[9]: history = model.fit(X_train, y_train, epochs=epochs, verbose=0)
      plt.plot(history.history['mse'])
      plt.plot(history.history['mae'])
      plt.show()
```



4 Teste do modelo criado

```
10]: minutes = ["00", "15", "30", "45"]

def plot_result(y_test, y_pred):
    n = N_STEPS_OUT//4
    x_data = []
    n = N_STEPS_OUT//4
    x_data = ["12h15", "12h30", "12h45"]
    for i in range(n-1):
        for j in range(4):
            x = str((i+13))+"h"+str(minutes[j])
            x_data.append(x)
    x_data.append("17h00")
    df = pd.DataFrame({"horas": x_data, "real": y_test, "predito": y_pred})

    data = pd.melt(df, id_vars=['horas'], value_vars=['real', "predito"],
        var_name='tipo', value_name='glicose (mg/dL)')

    plt.figure(figsize=(16,6))
```

```

g = sns.lineplot(data=data, x="horas", y="glicose (mg/dL)", hue="tipo",
↳markers=True, style="tipo", linewidth = 5)
g.set_xlabel("Horas", fontsize = 30)
g.set_ylabel("Glicose (mg/dL)", fontsize =30)
g.set_ylim([100, 240])
plt.setp(g.get_legend().get_texts(), fontsize='20')
plt.setp(g.get_legend().get_title(), fontsize='20')
plt.title(label = "Glicose real e predita ao longo das horas", fontsize =
↳20)
plt.savefig('imgs/resultado_lstm_individual_nao_diabetico.png')
plt.show()

def regression_report(y_true, y_pred):

    error = y_true - y_pred

    metrics = [
        ('mean absolute error', mean_absolute_error(y_true, y_pred)),
        ('mean squared error', mean_squared_error(y_true, y_pred)),
        ('root mean squared error', mean_squared_error(y_true, y_pred)**0.5)
    ]

    print('Metrics for regression:')
    for metric_name, metric_value in metrics:
        print(f'{metric_name:>25s}: {metric_value: >20.3f} {(1 -
↳metric_value)*100: >20.3f}%')

```

```

[11]: y_pred = model.predict(X_test)
      regression_report(y_test.reshape(y_pred.shape[0], y_pred.shape[1]), y_pred)

```

1/1 [=====] - 1s 987ms/step

Metrics for regression:

mean absolute error:	0.233	76.656%
mean squared error:	0.077	92.337%
root mean squared error:	0.277	72.318%

```

[12]: y_test_ = scaler.inverse_transform(np.mean(y_test, axis=0))
      y_pred_ = scaler.inverse_transform(np.mean(y_pred, axis=0).reshape(-1, 1))
      plot_result(y_test_.reshape(1, -1)[0], y_pred_.reshape(1, -1)[0])

```

**APÊNDICE G – PARECER CONSUBSTANCIADO DO CEP****PARECER CONSUBSTANCIADO DO CEP****DADOS DO PROJETO DE PESQUISA**

Título da Pesquisa: Modelo preditivo do nível glicêmico por monitoramento flash em indivíduos portadores de diabetes mellitus tipo II

Pesquisador: MARCELO VASCONCELOS

Área Temática:

Versão: 2

CAAE: 40854120.5.0000.5494

Instituição Proponente: UNIVERSIDADE BRASIL

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.506.553

Apresentação do Projeto:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020. Ver campo "Conclusões ou Pendências e Lista de Inadequações".

Objetivo da Pesquisa:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020. Ver campo "Conclusões ou Pendências e Lista de Inadequações".

Avaliação dos Riscos e Benefícios:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020. Ver campo "Conclusões ou Pendências e Lista de Inadequações".

Comentários e Considerações sobre a Pesquisa:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020. Ver campo "Conclusões ou Pendências e Lista de Inadequações".

Considerações sobre os Termos de apresentação obrigatória:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020. Ver campo "Conclusões ou Pendências e Lista de Inadequações".

Conclusões ou Pendências e Lista de Inadequações:

Trata-se de análise de resposta ao parecer pendente número 4.474.266 de parecer emitido em 18/12/2020:

Endereço: RUA CAROLINA FONSECA, 235

Bairro: ITAQUERA

UF: SP

Município: SAO PAULO

CEP: 08.230-030

Telefone: (11)2052-0461

E-mail: comite.etica.sp@universidadebrasil.edu.br



Continuação do Parecer: 4.506.553

1 - TCLE precisa de uma revisão conforme sub-itens abaixo:

- a) O texto aparece quebrado por muitos parênteses no meio do texto, atrapalhando o entendimento adequado pelo participante.
- b) Na exposição dos riscos aparece a frase a seguir que está em desacordo com os riscos apresentados no projeto: "A realização deste estudo não irá expor você a qualquer tipo de risco ou constrangimento." Como o pesquisador por garantir a ausência total de risco e constrangimento? Recomendamos retirar esta frase e adequar os riscos conforme descrito no projeto.
- c) O termo precisa ser revisado, evitando quebras de frase que impossibilitam o entendimento, como por exemplo: "Os pesquisadores envolvidos com o referido projeto são (O estudo está " ou "Recebi, por outro lado, os esclarecimentos sobre os possíveis desconfortos e riscos decorrentes do estudo como: (A realização deste..." ou ainda "CIDADE, São Luís - Ma 05 de novembro de 2020"
- d) Corrigir o telefone do CEP (11) 2052-0461
- e) Quanto a indenização, acrescentar: "Fui informado(a) também, que conforme as Resoluções do Conselho Nacional de Saúde serão aplicadas as garantias previstas de assistência integral, gratuita e pelo tempo que for necessário e direito de solicitação de indenização por danos decorrentes da pesquisa, cujo mérito será julgado em sua instância adequada, conforme a Resolução CNS nº 466 de 2012, itens III.2.o, IV.3.c, V.6 e V.7."
- f) Quanto aos custos ao participante, acrescentar: "Fui informado que não terei gastos para participar desta pesquisa; meu transporte e alimentação (e do acompanhante), se necessário, serão de responsabilidade dos pesquisadores."
- g) Quanto as assinaturas do TCLE: "Caso eu aceite participar, rubricarei cada página deste termo, apresentado em duas vias, e assinarei ao final juntamente com o pesquisador responsável ou pelas pessoas por ele delegadas, indicando meu consentimento em participar da pesquisa (Resolução CNS n 466 de 2012, item IV.5.d), das quais uma via me será entregue."

RESPOSTA:

- a) foi retirado os parênteses para melhor entendimento do texto.
- b) A sentença foi alterada para: algum tipo de alergia ao sensor após aplicação, desconforto ao dormir com a sensor, dificuldade do manuseio do leitor e do aplicativo de contagem de carboidratos, desconforto a responder o questionário relacionado ao tema da pesquisa.
- c) A sentença foi alterada para: os pesquisadores envolvidos com o referido projeto são, o doutorando Marcelo Henrique de Vasconcelos Mourão,

Endereço: RUA CAROLINA FONSECA, 235

Bairro: ITAQUERA

UF: SP

Município: SAO PAULO

CEP: 08.230-030

Telefone: (11)2052-0461

E-mail: comite.etica.sp@universidadebrasil.edu.br



Continuação do Parecer: 4.506.553

sob a orientação do Prof. Dr. Marcello Magri Amaral. contato pelos telefones, doutorando (98) 98135-6312, orientador (11) 99780-1966. O estudo está vinculado ao Programa de Pós-Graduação em Engenharia Biomédica, da Universidade Brasil Campus Itaquera SP, sendo assegurada a assistência durante toda pesquisa, bem como me é garantido o livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo.

d) (11) 2052-0461.

e) "Fui informado(a) também, que conforme as Resoluções do Conselho Nacional de Saúde serão aplicadas as garantias previstas de assistência integral, gratuita e pelo tempo que for necessário e direito de solicitação de indenização por danos decorrentes da pesquisa, cujo mérito será julgado em sua instância adequada, conforme a Resolução CNS nº 466 de 2012, itens III.2.o, IV.3.c, V.6 e V.7."

f) "Fui informado que não terei gastos para participar desta pesquisa; meu transporte e alimentação (e do acompanhante), se necessário, serão de responsabilidade dos pesquisadores."

g) "Caso eu aceite participar, rubricarei cada página deste termo, apresentado em duas vias, e assinarei ao final juntamente com o pesquisador responsável ou pelas pessoas por ele delegadas, indicando meu consentimento em participar da pesquisa (Resolução CNS n 466 de 2012, item IV.5.d), das quais uma via me será entregue."

ANÁLISE:

Atendida.

Considerações Finais a critério do CEP:

Ressalta-se que cabe ao pesquisador responsável encaminhar os relatórios parciais e final da pesquisa, por meio da Plataforma Brasil, via notificação do tipo "relatório" para que sejam devidamente apreciadas no CEP, conforme Norma Operacional CNS nº 001/13, item XI.2.d.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações	PB_INFORMAÇÕES_BÁSICAS_DO_P	06/01/2021		Aceito

Endereço: RUA CAROLINA FONSECA, 235

Bairro: ITAQUERA

UF: SP

Município: SAO PAULO

CEP: 08.230-030

Telefone: (11)2052-0461

E-mail: comite.etica.sp@universidadebrasil.edu.br



Continuação do Parecer: 4.506.553

UNIVERSIDADE BRASIL



Básicas do Projeto	ETO_1603169.pdf	15:41:21		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEREVISADO.pdf	06/01/2021 15:40:53	MARCELO VASCONCELOS	Aceito
Outros	CARTARESPOSTA.pdf	06/01/2021 15:40:21	MARCELO VASCONCELOS	Aceito
Outros	lattesalunomarcelo.pdf	04/12/2020 09:53:43	MARCELO VASCONCELOS	Aceito
Outros	Lattesprofmarcellol.pdf	04/12/2020 09:52:33	MARCELO VASCONCELOS	Aceito
Declaração de concordância	declaracao.pdf	11/11/2020 21:15:34	MARCELO VASCONCELOS	Aceito
Projeto Detalhado / Brochura Investigador	PROJETO_pdf.pdf	10/11/2020 16:51:09	MARCELO VASCONCELOS	Aceito
Folha de Rosto	Folha_de_Rosto.pdf	10/11/2020 16:35:54	MARCELO VASCONCELOS	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	05/11/2020 15:02:19	MARCELO VASCONCELOS	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO PAULO, 21 de Janeiro de 2021

Assinado por:

**DANIEL SOUZA FERREIRA
MAGALHAES**

Endereço: RUA CAROLINA FONSECA, 235

Bairro: ITAQUERA

UF: SP

Município: SAO PAULO

CEP: 08.230-030

Telefone: (11)2052-0461

E-mail: comite.etica.sp@universidadebrasil.edu.br

Endereço: RUA CAROLINA FONSECA, 235

Bairro: ITAQUERA

UF: SP

Município: SAO PAULO

CEP: 08.230-030

Telefone: (11)2052-0461

E-mail: comite.etica.sp@universidadebrasil.edu.br