

Classificação de Níveis de Obesidade através de Redes Neurais Multicamadas: Uma Abordagem Baseada em Hábitos de Vida e Dados Antropométricos

1º Miqueias Nilton de Araujo

Curso de Engenharia de Computação
Universidade Federal Rural do Semi-Árido (UFERSA)
Rio Grande do Norte, Brasil
miqueias.araujo@alunos.ufersa.edu.br

2º Rosana C. B. Rego

Departamento de Engenharias e Tecnologia
Universidade Federal Rural do Semi-Árido (UFERSA)
Rio Grande do Norte, Brasil
rosana.rego@ufersa.edu.br

Abstract—Obesity and the risk of cardiovascular disease represent growing challenges for global public health. Preventive diagnosis and accurate classification of risk levels are fundamental for effective medical interventions. This study presents an artificial intelligence-based approach, using a multilayer perceptron neural network to classify the obesity level of individuals based on their eating habits and physical condition. The proposed model, structured with dense layers, ReLU activation function, dropout regularization, and Adam optimizer, achieved an overall accuracy of 94% on the test data with consistent generalization. Metric analysis demonstrated remarkable robustness, achieving an F1-Score of 1.00 for the Grade III Obesity class, with zero false negatives in the highest risk category.

Index Terms—Artificial Intelligence, Artificial Neural Networks, Multilayer Perceptron, Obesity, Public Health, Decision Support Systems.

I. INTRODUÇÃO

A transição nutricional global resultou em um aumento alarmante da prevalência da obesidade, fator de risco central para patologias como diabetes mellitus e apneia do sono [10] [12]. A complexidade dos fatores que determinam o índice de massa corporal, incluindo genética, comportamento alimentar e atividade física, torna a classificação manual uma tarefa suscetível a erros [12] [15]. A prevalência global da obesidade mais do que dobrou nas últimas quatro décadas, e hoje afeta mais de um bilhão de pessoas em todo o mundo. Reconhecida como uma condição ligada a muitas doenças crônicas, a obesidade impacta diretamente na qualidade de vida e na redução da expectativa de vida [16].

Em 2021, a prevalência global da doença cardiovascular (DCV) atingiu 612 milhões de pessoas, representando 26,8% de todas as mortes mundiais, com um crescimento de 0,88% ao longo das últimas três décadas [14]. Notavelmente, 79,5% de todos os anos de vida saudável perdidos ajustados às incapacidades foram atribuídos a 11 fatores de risco, sendo o índice de massa corporal (IMC) o que apresentou a maior associação [14]. Não obstante, estudos de prevalência mostram que dois terços das mortes relacionadas à obesidade são atribuíveis

à DCV [6], [14]. Particularmente no Brasil, dados de 2025 sugerem que 68% dos adultos possuem IMC ≥ 25 kg/m², e que 31% dessa população esteja vivendo com a obesidade. Em 2021, houve 60.913 mortes prematuras associadas ao IMC elevado no país [16].

A obesidade também é um dos principais determinantes da DCV. Estudos epidemiológicos prospectivos demonstram que a obesidade está associada a um risco elevado de eventos de doença arterial coronariana (DAC) e de mortalidade cardiovascular (CV) [9]. Sua influência ocorre de forma indireta, por meio do aumento de fatores de risco CV tradicionais como diabetes tipo 2 (DM2), dislipidemia (DLP) e hipertensão arterial sistêmica (HAS), mas também por efeito direto do estado inflamatório induzido pela adiposidade na estrutura e função cardíacas [2], [3].

Historicamente, a abordagem terapêutica para a obesidade tem se centrado em intervenções relacionadas ao estilo de vida e opções farmacológicas com eficácia limitada. Contudo, a introdução da nova geração de medicações antiobesidade possibilitou a obtenção de perdas de peso mais significativas e sustentáveis. Ademais, com a complexidade crescente das opções terapêuticas disponíveis e os benefícios observados em desfechos ligados à síndrome cardiorrenal metabólica; tornou-se evidente a necessidade de novas ferramentas de estratificação que orientem a seleção do tratamento para cada situação clínica [4], [5].

Este artigo apresenta a aplicação de uma rede neural para a classificação de níveis de obesidade, detalhando as etapas de engenharia de atributos e a análise de desempenho por meio de matrizes de confusão e métricas de precisão e revocação.

II. LEVANTAMENTO BIBLIOGRÁFICO

Com a disseminação da Internet, as tecnologias da informação e comunicação (TICs) impulsionaram o avanço de produtos e serviços, sobretudo aqueles que funcionam de forma integrada. No setor da saúde, isso se refletiu em serviços mais eficientes e aprimorados, conforme destacado

por Rocha et al. (2016) [11]. Essa área de estudo é conhecida como mobileHealth (ou mHealth), sendo que a World Health Organization (WHO, 2011) a definiu como “as práticas médicas e de saúde pública suportadas por dispositivos móveis, como telefones celulares, dispositivos de monitoramento de pacientes, assistentes digitais e outros dispositivos sem fio”. Syed et al. (2019) [13] indicam que a Internet das Coisas Médicas, ou Internet of Medical Things (IoMT), representa uma inovação significativa na indústria da saúde, pois tem a capacidade de conectar profissionais de saúde, cuidadores, pacientes, dados e sistemas, visando oferecer serviços de saúde mais inteligentes. Cada vez mais, métodos de inteligência artificial estão sendo explorados e implementados no âmbito da saúde. Olivera et al. (2017) [7] aplicaram algoritmos de aprendizado de máquina para identificar diabetes não diagnosticada. Bejnordi et al. (2017) [1] demonstraram que algoritmos podem interpretar imagens de câncer de mama, alcançando desempenhos comparáveis aos de profissionais humanos na análise.

Safaei et al. (2021) [12] publicaram uma revisão de literatura bastante completa, citando as causas e consequências da obesidade e ainda fizeram um levantamento entre vários estudos, sobre as técnicas de machine learning (ML) para a predição da obesidade. Dentre os artigos analisados, Multi-layer Perceptron Feed Forward Artificial Neural Networks (MLFFNN), K-Nearest Neighbors (K-NN) e Support Vector Machines (SVM) foram os algoritmos mais utilizados, sendo o MLFFNN o que apresentou melhores resultados. Este trabalho mostrou que as causas são diversas, e incluem fatores que, normalmente, não são analisados, como por exemplo, a situação socioeconômica, fatores demográficos ou fatores psicológicos.

Rezende et al. (2021) [10] publicaram um artigo mostrando como o excesso de peso e a obesidade estão diretamente relacionadas com o aumento das Doenças Crônicas não Transmissíveis (DCNT) e os impactos gerados no Sistema Único de Saúde (SUS). Em 2019, 56,1% das mortes causadas por estas doenças foram em adultos de 30 a 69 anos, sendo consideradas mortes prematuras. Ainda neste estudo, dados apontam que em 2030, 68% da população brasileira estará acima do peso e destacam como a obesidade necessita de um olhar crítico e que políticas públicas precisam ser criadas para combater essa epidemia. Não basta acreditar que somente fatores individuais, como a genética, são os responsáveis por esse aumento. Citam ainda o crescimento acelerado de pessoas acima do peso nos EUA após a década de 70 e listam as mudanças nas leis agrícolas no país, o marketing acelerado, a disponibilidade e acesso a alimentos com alto teor calórico, e a introdução no sistema alimentar de adoçantes como xarope de milho e com alto teor de frutose como fatores que alteraram o padrão alimentar de toda a população. Além da preocupação com a saúde da população, os gastos com internações ou complicações médicas decorrentes do sobrepeso são alarmantes. Em 2019, houve um gasto de 6,8 bilhões de reais com as DCNTs somente no Brasil, sendo que 22% desse valor foi por fatores de sobrepeso e obesidade, de acordo com a pesquisa de Rezende. Assim sendo, se for possível fazer um

levantamento da porcentagem da população que está em risco de ficar obesa, e consequentemente, desenvolver as DCNTs, ações mais eficazes poderão ser tomadas.

III. MODELO DE REDE NEURAL PROPOSTO

A. Detalhamento do Problema

A obesidade e o sobrepeso são fatores de risco primários para o desenvolvimento de Doenças Cardiovasculares (DCV), diabetes e outras síndromes metabólicas. O diagnóstico clínico tradicional baseia-se frequentemente apenas no Índice de Massa Corporal (IMC), que pode ser redutivo por não considerar a complexidade dos hábitos de vida do paciente.

Do ponto de vista computacional, o problema formula-se como uma tarefa de classificação supervisionada multiclasse. Seja $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ o conjunto de dados com $N = 2111$ amostras, onde $\mathbf{x}_i \in \mathbb{R}^d$ é o vetor de características de entrada, com $d = 17$ atributos iniciais (após pré-processamento, a dimensionalidade aumenta para $d' = 26$ devido à codificação one-hot), e $y_i \in \mathcal{Y} = \{0, 1, 2, 3, 4, 5, 6\}$ representa a classe alvo correspondente ao nível de obesidade.

O objetivo é aprender uma função $f : \mathbb{R}^{d'} \rightarrow \mathcal{Y}$ que minimize o erro de classificação esperado:

$$f^* = \arg \min_{f \in \mathcal{F}} \mathbb{E}_{(\mathbf{x}, y) \sim \mathcal{D}} [\mathcal{L}(f(\mathbf{x}), y)] \quad (1)$$

onde $\mathcal{F}$ é o espaço de hipóteses (arquiteturas MLP) e $\mathcal{L}$ é a função de perda (entropia cruzada).

O desafio consiste em desenvolver um modelo capaz de capturar padrões não lineares complexos que mapeiem o vetor de características multidimensional $\mathbf{x}$ para a variável alvo y , mitigando erros de classificação, especialmente nas categorias de maior severidade.

B. Dados Utilizados

Os dados consistem na estimativa dos níveis de obesidade em pessoas do México, Peru e Colômbia, com idades entre 14 e 61 anos e diversos hábitos alimentares e condições físicas. Os dados foram coletados por meio de uma plataforma online com um questionário onde usuários anônimos responderam a cada pergunta. As informações foram então processadas, obtendo-se 17 atributos e 2111 registros [8].

O conjunto de dados possui tamanho total de $N = 2111$ amostras, com dimensão inicial de $d = 17$ atributos (16 preditores e 1 variável alvo). Após o pré-processamento, a dimensionalidade aumenta para $d' = 26$ características devido à codificação one-hot encoding de variáveis categóricas. O problema envolve $K = 7$ classes distintas. A divisão dos dados seguiu a proporção padrão $\mathcal{D}_{\text{train}} \cup \mathcal{D}_{\text{test}} = \mathcal{D}$, com $|\mathcal{D}_{\text{train}}| = 0.8N = 1689$ amostras para treinamento e $|\mathcal{D}_{\text{test}}| = 0.2N = 422$ amostras para teste.

Os atributos estão estabelecidos nas seguintes categorias: Variáveis Antropométricas e Pessoais: Gênero, Idade, Altura, Peso e Histórico familiar de sobrepeso. Hábitos Alimentares: Consumo frequente de alimentos altamente calóricos (FAVC), Frequência de consumo de vegetais (FCVC), Número de refeições principais (NCP), Consumo de alimentos entre as

refeições (CAEC), Consumo diário de água (CH2O) e Monitoramento de calorias ingeridas (SCC). Hábitos de Estilo de Vida: Fumo (SMOKE), Frequência de atividade física (FAF), Tempo de uso de dispositivos tecnológicos (TUE), Consumo de álcool (CALC) e Meio de transporte principal (MTRANS). Variável Alvo (Target): NObesy, dividida em 7 classes clínicas: Insufficient Weight (Abaixo do peso), Normal Weight (Peso normal), Overweight Level I (Sobrepeso Grau I), Overweight Level II (Sobrepeso Grau II), Obesity Type I (Obesidade Grau I), Obesity Type II (Obesidade Grau II), Obesity Type III (Obesidade Grau III) [8].

C. Pré-processamento dos Dados

O pipeline de pré-processamento estruturou-se em três etapas:

Codificação da Variável Alvo (Label Encoding): As sete categorias textuais da saída y foram mapeadas para um vetor de inteiros por meio de uma função bijetora $\phi: \mathcal{Y}_{\text{text}} \rightarrow \mathcal{Y} = \{0, 1, 2, 3, 4, 5, 6\}$:

$$\phi(\text{classe}) = \begin{cases} 0, & \text{Peso Insuficiente} \\ 1, & \text{Peso Normal} \\ 2, & \text{Sobrepeso Grau I} \\ 3, & \text{Sobrepeso Grau II} \\ 4, & \text{Obesidade Tipo I} \\ 5, & \text{Obesidade Tipo II} \\ 6, & \text{Obesidade Tipo III} \end{cases} \quad (2)$$

Transformação de Variáveis Categóricas (One-Hot Encoding): Atributos categóricos de entrada foram transformados em matrizes binárias ortogonais (ausentes de hierarquia). Uma variável categórica x_j com k categorias distintas $\{c_1, c_2, \dots, c_k\}$ foi transformada em k variáveis binárias (dummy variables) $\{x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(k)}\}$, onde:

$$x_j^{(i)} = \begin{cases} 1, & \text{se } x_j = c_i \\ 0, & \text{caso contrário} \end{cases} \quad (3)$$

Padronização Estatística (Z-Score Normalization): Para evitar que atributos com grandes amplitudes dominem o processo de otimização, aplicou-se o método de escalonamento padrão (Standard Scaler). Seja $\mathbf{x}_j = (x_{1j}, x_{2j}, \dots, x_{Nj})^T$ o vetor contendo todos os valores da característica j . A transformação Z-score é dada por:

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (4)$$

onde:

$$\mu_j = \frac{1}{N} \sum_{i=1}^N x_{ij}, \quad \sigma_j = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{ij} - \mu_j)^2} \quad (5)$$

Após a transformação, cada característica numérica possui média $\mu = 0$ e desvio padrão $\sigma = 1$, garantindo que todas as features contribuam equitativamente para a atualização dos pesos durante o treinamento via gradiente descendente.

D. Formulação Matemática do Modelo

A arquitetura selecionada para resolver este problema foi um MLP do tipo Feedforward totalmente conectado. A topologia da rede é especificada formalmente como:

$$\begin{aligned} \mathbb{R}^{26} &\rightarrow \text{Dense}(64) \rightarrow \text{Dropout}(0.2) \\ &\rightarrow \text{Dense}(32) \rightarrow \text{Dropout}(0.2) \\ &\rightarrow \text{Dense}(7) \rightarrow \mathbb{R}^7. \end{aligned} \quad (6)$$

onde a camada de entrada recebe vetores $\mathbf{x} \in \mathbb{R}^{26}$, a primeira camada oculta possui 64 neurônios com ativação ReLU seguida de Dropout ($p = 0.2$), a segunda camada oculta contém 32 neurônios com ativação ReLU seguida de Dropout ($p = 0.2$), e a camada de saída apresenta 7 neurônios com ativação Softmax. O modelo possui um total de $(26 \times 64 + 64) + (64 \times 32 + 32) + (32 \times 7 + 7) = 3943$ parâmetros treináveis.

O mapeamento do vetor de entrada $\mathbf{x}$ para a predição de saída $\hat{y}$ ocorre por meio de sucessivas transformações lineares seguidas de ativações não-lineares. Para uma camada oculta arbitrária l , a saída dos neurônios $\mathbf{h}^{(l)}$ é computada por:

$$\mathbf{h}^{(l)} = f(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}) \quad (7)$$

Onde $\mathbf{W}^{(l)} \in \mathbb{R}^{n_l \times n_{l-1}}$ é a matriz de pesos que conecta a camada $l-1$ à camada l , $\mathbf{b}^{(l)} \in \mathbb{R}^{n_l}$ é o vetor de vies (bias), $f(\cdot)$ é a função de ativação não-linear, e $\mathbf{h}^{(0)} = \mathbf{x}$ representa a entrada da rede.

Nas duas camadas ocultas do modelo (com 64 e 32 neurônios), adotou-se a função ReLU (Rectified Linear Unit), que mitiga o problema do desvanecimento do gradiente e introduz esparsidade nas representações:

$$f_{\text{ReLU}}(x) = \max(0, x) = \begin{cases} x, & \text{se } x > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (8)$$

Na camada de saída (com 7 neurônios, correspondentes às classes $\mathcal{Y} = \{0, 1, 2, 3, 4, 5, 6\}$), aplicou-se a função Softmax, que converte os valores brutos da rede ($\mathbf{z}$) em uma distribuição de probabilidade condicional para cada classe i :

$$P(y = i|\mathbf{x}) = \sigma_i(\mathbf{z}) = \frac{e^{z_i}}{\sum_{j=0}^6 e^{z_j}}, \quad \text{com } \sum_{i=0}^6 P(y = i|\mathbf{x}) = 1 \quad (9)$$

O ajuste da matriz de pesos $\mathbf{W}$ foi realizado iterativamente pelo otimizador Adam (Adaptive Moment Estimation) [17], que combina momento com taxa de aprendizado adaptativa. O algoritmo buscou minimizar o erro do modelo em relação aos dados reais. O erro foi quantificado matematicamente pela função de Entropia Cruzada Categórica Esparsa (Sparse Categorical Crossentropy):

$$\mathcal{L}(\theta) = -\frac{1}{|\mathcal{B}|} \sum_{n \in \mathcal{B}} \log P(y_n|\mathbf{x}_n; \theta) \quad (10)$$

Onde $\mathcal{B}$ é um mini-batch de treinamento, $|\mathcal{B}|$ é o tamanho do batch, e $\theta = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}_{l=1}^L$ representa todos os parâmetros treináveis da rede.

Para prevenir overfitting e melhorar a capacidade de generalização, aplicou-se a técnica de Dropout com taxa $p = 0.2$ após cada camada oculta. Durante o treinamento, cada neurônio é desativado com probabilidade p :

$$\mathbf{h}_{\text{dropout}}^{(l)} = \mathbf{m} \odot \mathbf{h}^{(l)}, \quad \text{onde } m_j \sim \text{Bernoulli}(1 - p) \quad (11)$$

onde $\odot$ denota multiplicação elemento a elemento e $\mathbf{m}$ é uma máscara binária.

A configuração completa dos hiperparâmetros utilizados no treinamento incluiu o otimizador Adam com taxa de aprendizado inicial $\alpha = 0.001$ e parâmetros $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-7}$. O treinamento utilizou mini-batches de tamanho $|\mathcal{B}| = 32$ amostras por um total de $E = 100$ épocas. A taxa de Dropout foi configurada em $p = 0.2$ para regularização, e os pesos foram inicializados usando o método Glorot Uniform. A função de perda empregada foi a Sparse Categorical Crossentropy, e a métrica de avaliação durante o treinamento foi a acurácia.

IV. RESULTADOS E SIMULAÇÕES

A. Avaliação e Métricas de Desempenho

O modelo MLP desenvolvido foi avaliado utilizando um conjunto de dados de teste $\mathcal{D}_{\text{test}}$, correspondente a 20% da amostra original ($|\mathcal{D}_{\text{test}}| = 422$ amostras), garantindo que a rede fosse testada com instâncias não vistas durante o treinamento.

Para avaliar o desempenho do modelo, utilizaram-se as métricas padrão de classificação: precisão (precision), revocação (recall) e F1-Score. Para uma classe i , estas métricas são definidas como:

$$\text{Precisão}_i = \frac{TP_i}{TP_i + FP_i}, \quad \text{Revocação}_i = \frac{TP_i}{TP_i + FN_i} \quad (12)$$

$$\text{F1-Score}_i = 2 \cdot \frac{\text{Precisão}_i \times \text{Revocação}_i}{\text{Precisão}_i + \text{Revocação}_i} = \frac{2 \cdot TP_i}{2 \cdot TP_i + FP_i + FN_i} \quad (13)$$

onde TP_i (verdadeiros positivos), FP_i (falsos positivos) e FN_i (falsos negativos) são computados a partir da matriz de confusão.

A acurácia global do classificador foi:

$$\text{Acurácia} = \frac{1}{|\mathcal{D}_{\text{test}}|} \sum_{i=1}^{|\mathcal{D}_{\text{test}}|} \mathbb{I}[\hat{y}_i = y_i] = 0.94 \quad (14)$$

evidenciando a capacidade da arquitetura de mapear os hábitos de vida e dados antropométricos para o nível de risco adequado em 94% dos casos.

Para uma análise mais granular do comportamento do modelo em um cenário multiclasse, extraiu-se o Relatório de Classificação (Tabela 1), que detalha as métricas de Precisão, Revocação e F1-Score para cada categoria.

Tabela 1
RELATÓRIO DE MÉTRICAS POR CLASSE

Classe de Risco	Precisão	Revocação	F1-Score
Peso Insuficiente	0.95	0.98	0.96
Peso Normal	0.92	0.89	0.90
Sobrepeso Grau I	0.86	0.89	0.88
Sobrepeso Grau II	0.94	0.92	0.93
Obesidade Tipo I	0.97	0.97	0.97
Obesidade Tipo II	0.98	0.97	0.97
Obesidade Tipo III	1.00	1.00	1.00

B. Análise dos Resultados via Matriz de Confusão

Gerou-se a Matriz de Confusão $\mathbf{C} \in \mathbb{N}^{7 \times 7}$, onde o elemento C_{ij} representa o número de instâncias da classe real i que foram classificadas como classe j . Formalmente:

$$C_{ij} = \sum_{k=1}^{|\mathcal{D}_{\text{test}}|} \mathbb{I}[y_k = i \wedge \hat{y}_k = j] \quad (15)$$

Os acertos correspondem à diagonal principal ($i = j$), enquanto os erros de classificação estão nos elementos fora da diagonal ($i \neq j$).

A Fig. 1 mostra a matriz de confusão obtida. Para a classe 6 de risco extremo (Obesidade Tipo III), o modelo atingiu uma precisão e revocação de $P_6 = R_6 = 1.00$, não registrando nenhum falso negativo ($FN_6 = 0$).

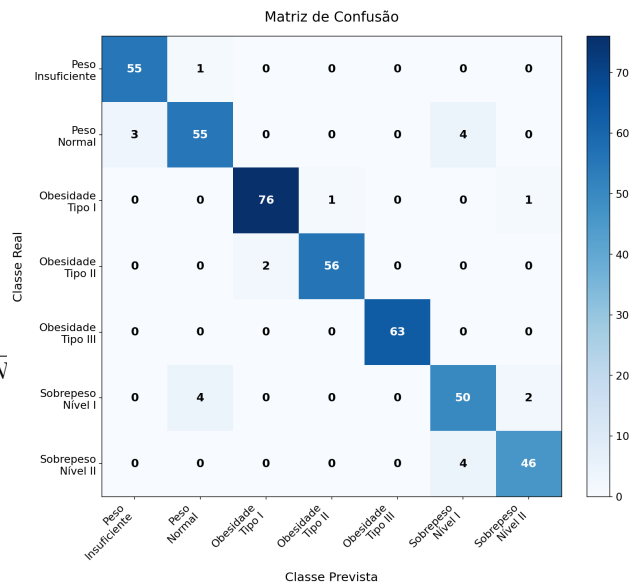


Fig. 1. Matriz de Confusão

Do ponto de vista clínico, a eliminação de falsos negativos nesta categoria é fundamental, pois garante que pacientes com necessidade de intervenção médica imediata não sejam subdiagnosticados. Um falso negativo nesta classe poderia resultar em ausência de tratamento adequado para indivíduos em alto risco cardiovascular.

Os desvios de predição (erros) concentraram-se majoritariamente entre classes adjacentes, particularmente na fronteira entre Peso Normal (classe 1) e Sobrepeso Grau I (classe 2). Esta sobreposição é compreensível dada a natureza contínua da biologia humana. Indivíduos próximos aos limiares de IMC frequentemente compartilham padrões comportamentais similares, tornando a separação matemática mais difícil para o algoritmo. Importante ressaltar que não foram observados erros de classificação entre classes distantes (por exemplo, Peso Normal classificado como Obesidade Tipo III), o que demonstra a capacidade do modelo de capturar a ordenação natural das classes.

V. CONCLUSÕES

O presente estudo demonstrou a aplicação de um MLP na classificação dos níveis de obesidade. A partir da análise de dados antropométricos e comportamentais, o modelo desenvolvido atingiu uma acurácia global de 94% no conjunto de teste, com F1-Score médio ponderado de 0.94, comprovando a correlação matemática não linear entre hábitos de vida diários e o estado nutricional dos indivíduos. O desempenho na detecção da classe de risco extremo (Obesidade Tipo III), com precisão e revocação de $P = R = 1.00$ e zero falsos negativos, reforça o potencial desta arquitetura como ferramenta de triagem para a prática clínica preventiva.

Parte significativa das variáveis do conjunto de dados do Kaggle provém de questionários preenchidos pelos próprios usuários. Dados autorrelatados sobre consumo de álcool, calorias e atividade física estão sujeitos ao viés de percepção e à subestimação comportamental. Os dados concentram-se nas populações do México, Peru e Colômbia. Consequentemente, o modelo pode apresentar uma queda de precisão se aplicado a populações de outros continentes com matrizes genéticas e dietas culturais diferentes.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Os autores declaram que utilizaram ferramentas de inteligência artificial, incluindo Google Gemini, exclusivamente para auxiliar na melhoria da linguagem e legibilidade deste manuscrito. O uso dessa ferramenta foi realizado sob supervisão humana, e os autores revisaram e editaram cuidadosamente o conteúdo, assumindo total responsabilidade pelo texto final publicado.

REFERENCES

- [1] BEJNORDI, B. E. et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*, v. 318, n. 22, p. 2199–2210, 12 2017. ISSN 0098-7484. Disponível em: <https://doi.org/10.1001/jama.2017.14585>.
- [2] Carmienie S, Freitag MH, Pischon T, Schlattmann P, Fankhaenel T, Goebel H, et al. General and abdominal obesity parameters and their combination in relation to mortality: a systematic review and meta-regression analysis. *Eur J Clin Nutr*. 2013;67(6):573–85.
- [3] Cerhan JR, Moore SC, Jacobs EJ, Kitahara CM, Rosenberg PS, Adami HO, et al. A pooled analysis of waist circumference and mortality in 650,000 adults. *Mayo Clin Proc*. 2014;89(3):335–45.
- [4] Godoy-Matos AF, Valério CM, Júnior WSS, de Araujo-Neto JM, Sposito AC, Suassuna JHR. CARDIAL-MS (CARDio-Renal-DIAbetes-Liver-Metabolic Syndrome): a new proposition for an integrated multisystem metabolic disease. *Diabetol Metab Syndr*. 2025 Jun 16;17(1):218. doi: 10.1186/s13098-025-01796-4. PMID: 40524210; PMCID: PMC12168269.
- [5] Khan SS, Matsushita K, Sang Y, et al. Development and Validation of the American Heart Association Predicting Risk of Cardiovascular Disease EVENTS (PREVENTTM) Equations. *Circulation* 2023.
- [6] Koskinas KC, Van Craenenbroeck EM, Antoniadis C, Blüher M, Gorter TM, Hanssen H, Marx N, McDonagh TA, Mingrone G, Rosengren A, Prescott EB; ESC Scientific Document Group. Obesity and cardiovascular disease: an ESC clinical consensus statement. *Eur Heart J*. 2024 Oct 7;45(38):4063–4098. doi:10.1093/eurheartj/ehae508. Erratum in: *Eur Heart J*. 2025 Mar 3;46(9):876. doi: 10.1093/eurheartj/ehae915. PMID: 39210706.
- [7] OLIVERA, A. R. et al. Comparison of machine-learning algorithms to build a predictive model for detecting undiagnosed diabetes-elsa-brasil: accuracy study. *Sao Paulo Medical Journal, SciELO Brasil*, v. 135, n. 3, p. 234–246, 2017.
- [8] Palechor, F. M., & de la Hoz Manotas, A. (2019). Dataset for estimation of obesity levels based on eating habits and physical condition in individuals from Colombia, Peru and Mexico. *Data in Brief*, 25, 104344. Kaggle Repository.
- [9] Powell-Wiley TM, Poirier P, Burke LE, Després JP, Gordon-Larsen P, Lavie CJ, Lear SA, Ndumele CE, Neeland IJ, Sanders P, St-Onge MP; American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; Council on Clinical Cardiology; Council on Epidemiology and Prevention; and Stroke Council. Obesity and Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Circulation*. 2021 May 25;143(21):e984–e1010. doi: 10.1161/CIR.0000000000000973. Epub 2021 Apr 22. PMID: 33882682.
- [10] REZENDE, L. F. M. et al. A EPIDEMIA DE OBESIDADE E AS DCNT. 2021. Disponível em: <https://rezendelfm.github.io/obesidade-e-as-dcnt/>. Acesso em: 03 abril 2026.
- [11] ROCHA, T. A. H. et al. Saúde móvel: novas perspectivas para a oferta de serviços em saúde. *Epidemiologia e Serviços de Saúde, SciELO Public Health*, v. 25, p. 159–170, 2016.
- [12] SAFARI, M. et al. A systematic literature review on obesity: Understanding the causes & consequences of obesity and reviewing various machine learning approaches used to predict obesity. *Computers in biology and medicine*, Elsevier, v. 136, p. 104754, 2021.
- [13] SYED, L. et al. Smart healthcare framework for ambient assisted living using iomt and big data analytics techniques. *Future Generation Computer Systems*, Elsevier, 2019.
- [14] Tan SCW, Zheng BB, Tang ML, Chu H, Zhao YT, Weng C. Global Burden of Cardiovascular Diseases and its Risk Factors, 1990–2021: A Systematic Analysis for the Global Burden of Disease Study 2021. *QJM*. 2025 Jan 23; hca022. doi:10.1093/qjmed/hca022. Epub ahead of print. PMID: 39847534.
- [15] WILBORN, C. et al. Obesity: Prevalence, Theories, Medical Consequences, Management, and Research Directions. *Journal of the International Society of Sports Nutrition*, v. 2, n. 2, p. 4–31, 2005. doi: 10.1186/1550-2783-2-2-4. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC2129146/>.
- [16] World Obesity Federation. *World Obesity Atlas 2025*. London: World Obesity Federation. 2025. Disponível em: <https://data.worldobesity.org/publications/?cat=23>
- [17] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” arXiv preprint arXiv:1412.6980, 2014.