

# Sistema Inteligente de Predição de Obesidade



Ferramenta de Apoio à Decisão Médica Baseada em  
Machine Learning

# A Obesidade como Desafio de Saúde Pública



## Contexto

A obesidade é uma condição médica multifatorial e crescente, atingindo pessoas de todas as idades e classes sociais. Exige intervenções precoces e estratégicas.



## Problema

Necessidade de uma ferramenta robusta e escalável para auxiliar médicos na triagem, diagnóstico preditivo e intervenção precoce do nível de obesidade.



## Objetivo

Desenvolver um modelo de Machine Learning com assertividade superior a 75% para classificar o nível de obesidade de pacientes.

# Solução Completa: Predição e Análise de Dados



## Aplicação de Predição

Interface intuitiva para diagnóstico em tempo real

- Inputs organizados em abas
- Cálculo de peso ideal
- Gráfico de probabilidades
- Feedback imediato ao paciente



## Painel Analítico

Dashboard estratégico com insights para a equipe médica

- Análise de fatores de risco
- Visualizações de dados
- Insights epidemiológicos
- Apoio à tomada de decisão

## Stack Tecnológico

Python • Streamlit • Machine Learning (Gradient Boosting Classifier) • Pandas • Scikit-learn

# Construção do Modelo Preditivo

## Dataset

Registros: 2,111

Arquivo: [obesity.csv](#)

## Estilo de Vida

### Comportamento

- SMOKE: Tabagismo
- CH2O: Consumo de água
- SCC: Monitoramento calórico
- FAF: Atividade física
- TER: Tempo em dispositivos
- CALC: Consumo de álcool
- MTRANS: Meio de transporte

## Modelo

### Gradient Boosting Classifier

Algoritmo robusto para classificação multiclasse, ideal para prever os 7 níveis de peso com alta precisão e interpretabilidade.

## Features Chave

### Hábitos Alimentares

- FAVC: Alimentos calóricos
- FCVC: Consumo de vegetais
- NCP: Refeições principais
- CAEC: Lanches entre refeições

## Demográficos

### Características Pessoais

- Gênero
- Idade
- Altura
- Peso
- Histórico familiar

## Performance

Assertividade Mínima

> 75%

# Transformando Dados em Ação Preventiva



## Inputs Organizados em Abas

Dados Pessoais, Hábitos Alimentares e Estilo de Vida para reduzir sobrecarga cognitiva



## Cálculo de Peso Ideal

Transforma o resultado em uma meta clara e açãoável para o paciente



## Gráfico de Probabilidades

Visualização da confiança do modelo em cada classe de classificação



## Recomendações Personalizadas

Orientações específicas baseadas no resultado da predição

## Visão de Negócio

*Ferramenta de triagem rápida e não invasiva para direcionar o paciente ao tratamento adequado*

The screenshot shows a web-based application titled "Sistema Inteligente de Predição de Obesidade". The main header includes a logo of a person with a scale, the title, and a subtitle: "Ferramenta de apoio à decisão médica baseada em Machine Learning". Below the header, there's a purple banner with the text "Entenda a Obesidade e seus Fatores de Risco". The main content area is divided into several sections: "Sobre o Sistema" (with a sub-section "O que é este sistema?"), "Como funciona?", "Por que é importante?", and "Fatores que Influenciam". The "Fatores que Influenciam" section lists "Dados Demográficos" (Idade, gênero, altura e peso), "Hábitos Alimentares" (Consumo de alimentos calóricos, Ingestão de vegetais e água, Frequência de refeições), and "Comportamentos". On the right side, there are social media sharing icons for LinkedIn, Facebook, and Twitter.

Interface da Aplicação de Predição

# Descobertas Chave para a Equipe Médica

Total de Pacientes

**2.111**

Média de Idade

**24.3<sup>a</sup>**

Média de IMC

**29.70**

% Obesidade (I, II, III)

**46%**



## Risco Genético Elevado

A proporção de obesidade é significativamente maior em pacientes com histórico familiar positivo, validando a importância do fator genético como preditor.

**Estratégia:** Reforçar o rastreamento precoce e a intervenção preventiva em famílias de risco.



## Atividade Física como Fator Crítico

A frequência de atividade física (FAF) é inversamente proporcional ao nível de peso, demonstrando seu papel crucial na prevenção e controle da obesidade.

**Estratégia:** Foco na promoção de exercícios como intervenção primária.

# Impacto e Evolução do Projeto

## Impacto Entregue

Solução completa que atende a todos os requisitos do Tech Challenge Fase 04, oferecendo valor real para a equipe médica através de:

- ✓ Modelo preditivo com assertividade acima de 75%
- ✓ Sistema de predição intuitivo e responsivo
- ✓ Dashboard analítico com insights estratégicos
- ✓ Deploy contínuo e acessibilidade 24/7

## Próximos Passos



### Integração com Prontuário

Conectar com sistemas de prontuário eletrônico via API



### Validação Clínica

Testar modelo com dados clínicos reais



### Recomendações Personalizadas

Módulo de dieta e exercícios customizado

# Performance do Modelo: Resultados Alcançados

## Validação Cruzada (5-Fold)

Acurácia Média: 95.74%

Folds: [95.51%, 95.50%, 96.68%, 95.26%, 95.73%]

## Teste em Holdout (20%)

Acurácia: 95.27%

Amostras: 423 pacientes

## Relatório de Classificação por Classe

Classe | Precisão | Recall | F1-Score | Suporte

Baixo Peso | 96% | 89% | 92% | 54

Peso Normal | 85% | 91% | 88% | 58

Sobrepeso I | 95% | 91% | 93% | 58

Sobrepeso II | 95% | 97% | 96% | 58

Obesidade I | 97% | 99% | 98% | 70

Obesidade II | 98% | 100% | 99% | 60

Obesidade III | 100% | 98% | 99% | 65

# Feature Engineering: Preparação dos Dados

---

## Pré-processamento

- ✓ StandardScaler: Normalização de features numéricas
- ✓ OneHotEncoder: Codificação de variáveis categóricas
- ✓ ColumnTransformer: Pipeline automático de transformações
- ✓ Stratified K-Fold: Validação cruzada com 5 folds
- ✓ Train-Test Split: 80% treino, 20% teste

## Features Numéricas

- Idade
- Altura
- Peso
- FCVC
- NCP
- Água/dia
- Atividade Física
- Tempo em Telas

## Features Categóricas

- Gênero
- Histórico Familiar
- FAVC
- Fuma
- Conta Calorias
- CAEC
- Álcool
- Transporte

## Algoritmo: Gradient Boosting Classifier (GBC)

Ensemble de árvores de decisão com boosting sequencial para classificação multiclasse com alta precisão

# Obrigado!

Acesse o Projeto



Aplicação de Predição

[predicao-obesidade-tech-4.streamlit.app](https://predicao-obesidade-tech-4.streamlit.app)



Painel Analítico

[painele-obesidade-tech-4.streamlit.app](https://painele-obesidade-tech-4.streamlit.app)



Repositório GitHub

[github.com/gustmacena/predicao-obesidade](https://github.com/gustmacena/predicao-obesidade)

Desenvolvido por

Gustavo Macena, Italo Costa, Raniel Richardeli, Lucas Siqueira e Alexandre Sluzek  
Tech Challenge - POSTECH Data Analytics