



8ª Feira Mineira de Iniciação Científica



De 09 a 29 de novembro de 2024

Ciências da Saúde  
**FEMIC JOVEM**

**Maria Fernanda Marques Nogueira**

**Carolina Beatriz Rosa Madógllo**  
**Mariane Machado Cardoso Luiz Pires.**

Doriane dos Santos Honorato Moura

**E E Gregoriano Canedo**  
**Monte Carmelo, MG Brasil**



[doriane.honorato@educacao.mg.gov](mailto:doriane.honorato@educacao.mg.gov)

## **Avaliação do desempenho de algoritmos de aprendizagem de máquina no diagnóstico da Diabetes Tipo 2**



# Apresentação

O projeto desenvolvido visa aplicar algoritmos de aprendizado de máquina para prever o risco de diabetes tipo 2 em mulheres, utilizando um conjunto de dados médicos fornecido pelo Instituto Nacional de Diabetes e Doenças Digestivas e Renais. A motivação deste estudo se dá pela crescente prevalência do diabetes tipo 2 e a importância de ferramentas preditivas no diagnóstico precoce.

A escolha do tema é justificada pela relevância da saúde pública, dado que o diabetes tipo 2 afeta milhões de pessoas globalmente, levando a complicações sérias. O uso de aprendizado de máquina tem o potencial de melhorar a precisão do diagnóstico e a intervenção clínica.



## Objetivos

- O objetivo principal deste trabalho é desenvolver e avaliar modelos de aprendizado de máquina para a predição da presença de diabetes tipo 2 em mulheres, utilizando dados médicos de pacientes. A proposta busca fornecer uma ferramenta de suporte ao diagnóstico médico, promovendo uma identificação precoce de pacientes com alto risco para diabetes tipo 2.
- **Objetivos Específicos:**
- Realizar a coleta e análise exploratória dos dados fornecidos pelo Instituto Nacional de Diabetes e Doenças Digestivas e Renais.
- Aplicar e comparar diferentes algoritmos de aprendizado de máquina, como Random Forest e Gradiente Descendente, para o desenvolvimento dos modelos preditivos.
- Avaliar o desempenho dos modelos utilizando métricas como acurácia, sensibilidade, especificidade e AUC.
- Propor um modelo Ensemble, combinando diferentes algoritmos, para obter maior precisão no diagnóstico de diabetes tipo 2.

# Metodologia



O projeto foi desenvolvido em etapas, iniciando com a coleta dos dados do Instituto Nacional de Diabetes e Doenças Digestivas e Renais, que incluía informações sobre mulheres indígenas Pima com idade superior a 21 anos. Uma análise exploratória foi conduzida para identificar padrões e entender as variáveis disponíveis no conjunto de dados.



A seguir, foram implementados algoritmos de aprendizado de máquina, como Random Forest e Gradiente Descendente, utilizando a linguagem Python na plataforma Google Colab. Cada modelo foi treinado e avaliado utilizando métricas como acurácia, sensibilidade e AUC. O processo também incluiu a aplicação de um modelo Ensemble, que combinou os algoritmos citados para melhorar a performance preditiva.

# Metodologia da Pesquisa

1

## Coleta de Dados

Obtenção do conjunto de dados do Instituto Nacional de Diabetes e Doenças Digestivas e Renais, contendo informações de mulheres indígenas Pima.

2

## Análise Exploratória

Realização de análise detalhada do banco de dados para compreender as variáveis e identificar padrões.

3

## Implementação de Algoritmos

Aplicação de algoritmos de aprendizado de máquina, como Random Forest e Gradiente Descendente, para construir modelos preditivos.

4

## Avaliação de Desempenho

Utilização de métricas como acurácia, sensibilidade, especificidade e AUC para avaliar o desempenho dos modelos.



# Resultados alcançados

- O projeto alcançou resultados significativos na predição da presença de diabetes tipo 2, com a implementação de algoritmos como Random Forest e Gradiente Descendente. O modelo Ensemble, que combina esses algoritmos, obteve uma acurácia superior a 86%, superando os modelos individuais. A sensibilidade e a especificidade apresentaram valores satisfatórios, demonstrando a capacidade dos modelos em identificar corretamente pacientes com e sem a doença.



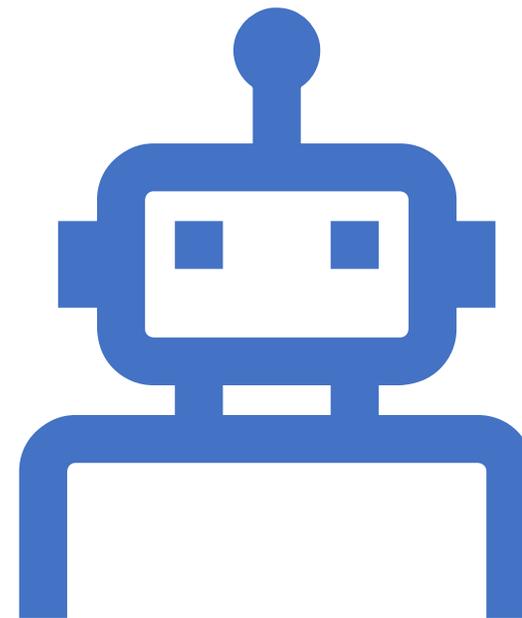
# Aplicabilidade dos resultados no cotidiano da sociedade

- O trabalho desenvolvido tem uma clara aplicabilidade na área de saúde pública, especialmente no diagnóstico precoce do diabetes tipo 2. O modelo preditivo proposto pode ser utilizado em ambientes clínicos para auxiliar médicos e profissionais de saúde a identificar pacientes em risco, permitindo intervenções mais rápidas e eficazes. Isso contribui diretamente para a prevenção de complicações graves relacionadas ao diabetes, como doenças cardiovasculares e renais.



# Criatividade e inovação

- Este projeto apresenta um caráter inovador ao utilizar algoritmos de aprendizado de máquina de forma combinada, através do modelo Ensemble, para melhorar a acurácia na predição do diabetes tipo 2. Ao explorar diferentes técnicas de machine learning, como Random Forest e Gradiente Descendente, e integrá-las em um modelo unificado, o projeto eleva o potencial preditivo em comparação com modelos tradicionais.

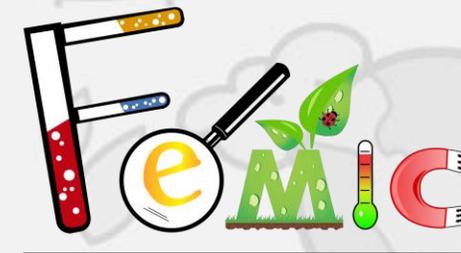


# Considerações finais

- Este projeto demonstrou a eficácia de algoritmos de aprendizado de máquina para a predição do diabetes tipo 2, destacando-se o modelo Ensemble, que apresentou excelente desempenho. Através da análise de dados e da construção de modelos preditivos, foi possível contribuir para a área da saúde pública, fornecendo uma ferramenta acessível e precisa para o diagnóstico precoce. O desenvolvimento deste trabalho reforça a importância da inovação tecnológica aplicada à saúde e à prevenção de doenças..



Gostaríamos de expressar nossa gratidão à Escola Estadual Gregoriano Canedo, à SRE Monte Carmelo, e ao ICEB - Núcleos de Pesquisa, cujas contribuições e apoio foram essenciais para a realização deste projeto.



7ª Feira Mineira de Iniciação Científica

**De 09 a 29 de novembro de 2024**

#### Realização



Associação Mineira de  
Pesquisa e Iniciação Científica



#### Apoiadores

UNIVERSIDADE  
DO ESTADO DE MINAS GERAIS | UEMG  
FACULDADE DE EDUCAÇÃO

MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÃO

GOVERNO FEDERAL  
**BRASIL**  
UNIÃO E RECONSTRUÇÃO

**CNPq**  
Conselho Nacional de Desenvolvimento  
Científico e Tecnológico

