

FIAP - Faculdade de Informática e Administração Paulista

DISRUPTIVE-ARCHITECTURES-IOT-IOB-GENERATIVE-IA

São Paulo

2024

ODONTO PREV

Integrantes:

RM553705 - Eduardo Rodrigues

RM553724 - Giovani Borba

RM553180 – Matheus Rodrigo da Silva

RESUMO

Este projeto utiliza aprendizado de máquina para prever a eficiência energética de edifícios com base em características como área construída e pontuação energética.

O objetivo é otimizar recursos e auxiliar na transição para práticas de construção mais sustentáveis.

1. Descrição do Dataset

O dataset utilizado contém as seguintes colunas:

- 10_Digit_BBL: Código de identificação do prédio.
- Street_Number: Número do endereço.
- DOF_Gross_Square_Footage: Área bruta construída (em pés quadrados).
- Energy_Efficiency_Grade: Classificação da eficiência energética.
- Energy_Star_1-100_Score: Pontuação de eficiência energética.

2. Estatísticas Descritivas do Dataset

Variável | Count | Mean | Std | Min | 25% | 50% | 75% | Max

10_Digit_BBL | 100 | 1.000952e+09 | 2.232645e+06 | 1.000050e+09 |

1.000160e+09 | 1.000220e+09 | 1.000290e+09 | 1.00762

Street_Number | 100 | 123.230 | 137.155 | 1.0 | 25.0 | 64.0 | 202.75 |

555.0

DOF_Gross_Square_Footage | 100 | 5.103363e+05 | 5.303829e+05 |
2.548300e+04 | 2.171008e+05 | 3.611830e+05 | 5.84440

Energy_Star_1-100_Score | 100 | 63.250 | 27.262 | 1.0 | 43.5 | 75.5 |
83.25 | 100.0

3. Pré-processamento dos Dados

Os dados passaram pelas seguintes etapas:

- Remoção de valores nulos.- Codificação da variável alvo (``Energy_Efficiency_Grade``) para valores numéricos.
- Normalização das variáveis de entrada (``DOF_Gross_Square_Footage`` e ``Energy_Star_1-100_Score``).

4. Modelos treinados

Foram utilizados os seguintes modelos de aprendizado de máquina:

1. **Random Forest** (modelo com melhor desempenho).
2. Gradient Boosting.
3. SVM.

5 - Modelos treinados

Foram utilizados os seguintes modelos de aprendizado de máquina:

1. **Random Forest** (modelo com melhor desempenho).
2. Gradient Boosting.
3. SVM.

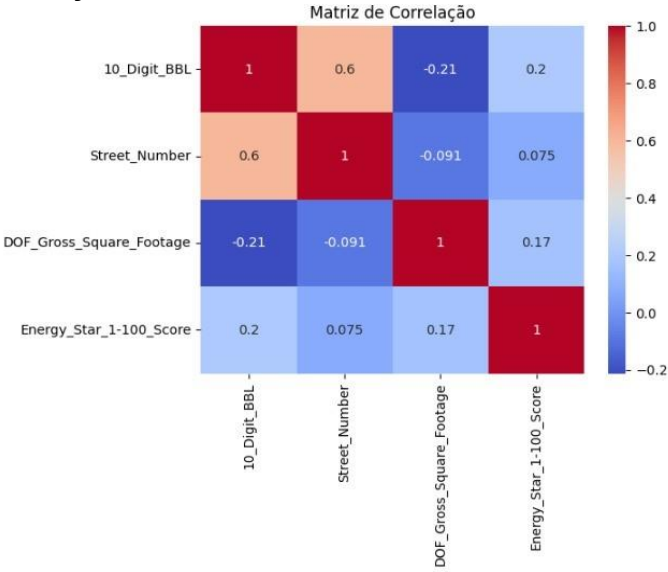
Resultados

O modelo Random Forest apresentou os melhores resultados, com as métricas a seguir:

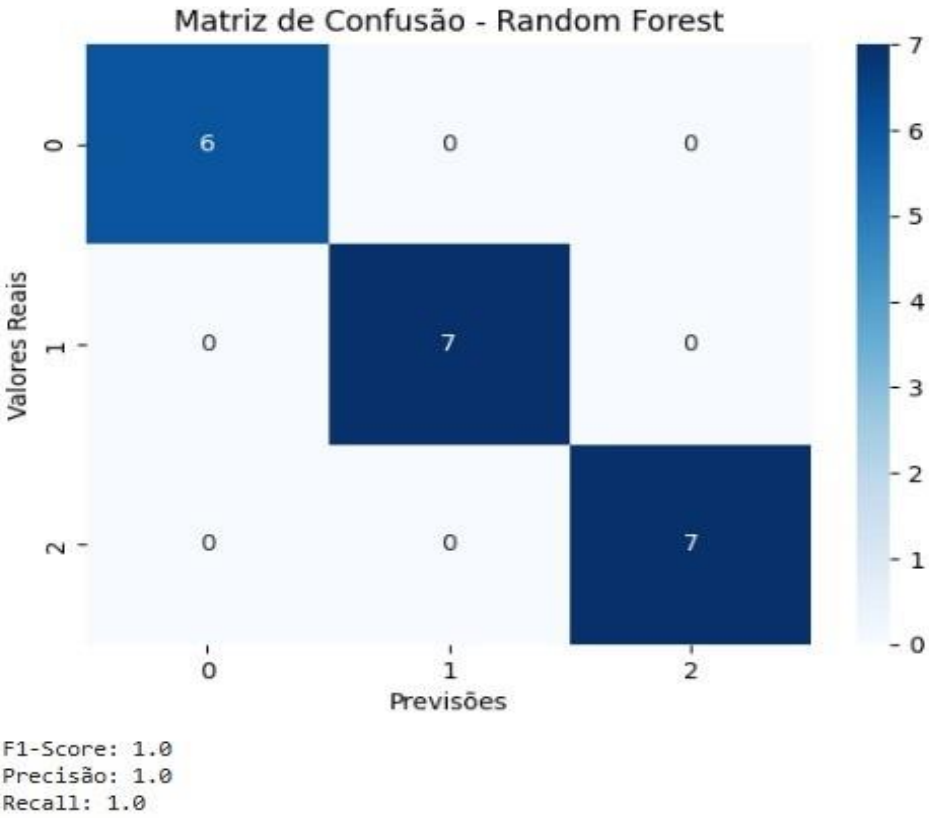
- **Acurácia:** 1.0
- **F1-Score:** 1.0
- **Precisão:** 1.0 - **Recall:** 1.0

A matriz de confusão demonstra que todas as previsões foram classificadas corretamente.

Matriz de Correlação:



Matriz de Confusão -



Conclusão

O modelo Random Forest demonstrou ser uma ferramenta eficiente para prever a eficiência energética de edifícios com

base nos dados fornecidos.

Com uma acurácia de 100%, o modelo é confiável para aplicações práticas. Como próximos passos, é possível

explorar mais variáveis e incorporar dados externos, como informações meteorológicas ou socioeconômicas.