

Previsão de Preço de Imóveis

Gustavo Navarro Felix



Como prever preço de imóveis?

O que é machine Learning?

Escolha de Variáveis

Pré-processamento

Escolha do Modelo

Visualização dos Resultados



Como prever preço de imóveis?

Prever o valor de um imóvel é como uma adivinhação. Você olha para suas características, seu tamanho, estilo, número de quartos e condição. E então diz qual o valor ele deve ter.

Graças aos avanços da tecnologia, podemos fazer isso em larga escala usando computadores.



O que é Machine Learning?

Colocamos computadores para analisar conjuntos de dados, eles então aprendem como fazer previsões ou tomar decisões utilizando esses dados como base, encontrando padrões e tendências.



Escolha de Variáveis

Para colocar o computadores para aprender com o conjunto de dados, precisamos garantir que seja entregue informações úteis e relevantes sobre o imóvel.

Correlação x Causalidade



Pré-processamento

Quando queremos fazer uma refeição é necessário lavar, descascar e cortar os ingredientes.

O pré-processamento segue a mesma ideia, para entregar os dados ao nosso modelo, precisamos preparar os dados:

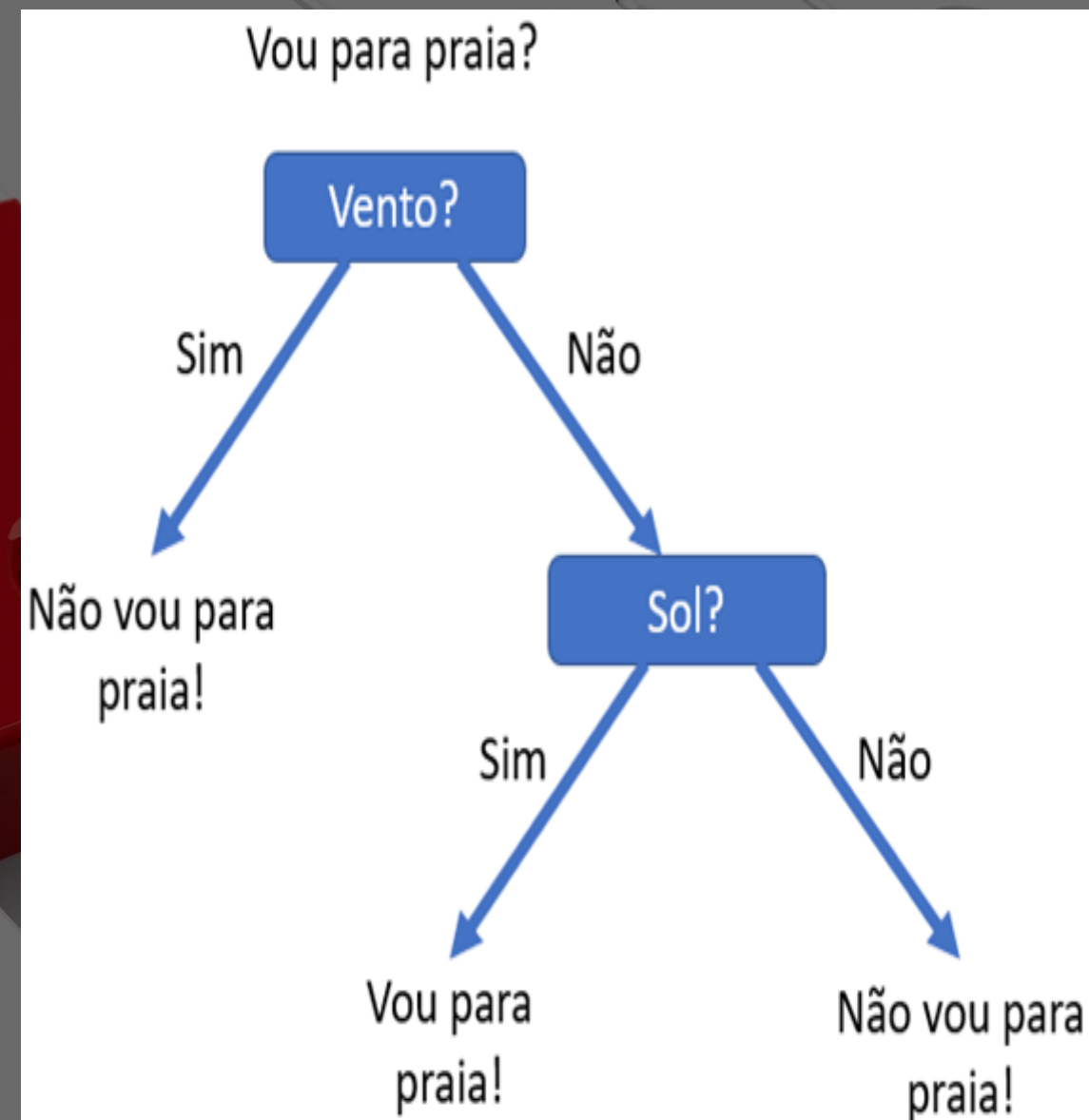
- Limpar os dados
- Remover valores discrepantes
- Remover valores desnecessários
- Tratar valores nulos.
- Organizar



Escolha do Modelo

É de extrema importância escolher a ferramenta certa para cada situação. Cada modelo possui um objetivo diferente.

Foi escolhido a Arvore de Decisão, esse modelo permite fazer previsões a partir de várias características, como o tamanho do terreno, localização e materiais.



Resultados

O treinamento do modelo foi feito inicialmente realizando testes, com os métodos Train-Test-Split e K-Fold Cross Validation:

Train-Test-Split: permite que o modelo seja treinado utilizando uma parte dos dados, e a outra parte é utilizada para testar e validar os resultados.

K-Fold Cross Validation: permite treinar o modelo várias vezes com diferentes partes dos dados, obtendo uma média dos resultados.

Train-Test-Split – Coeficiente de Determinação:

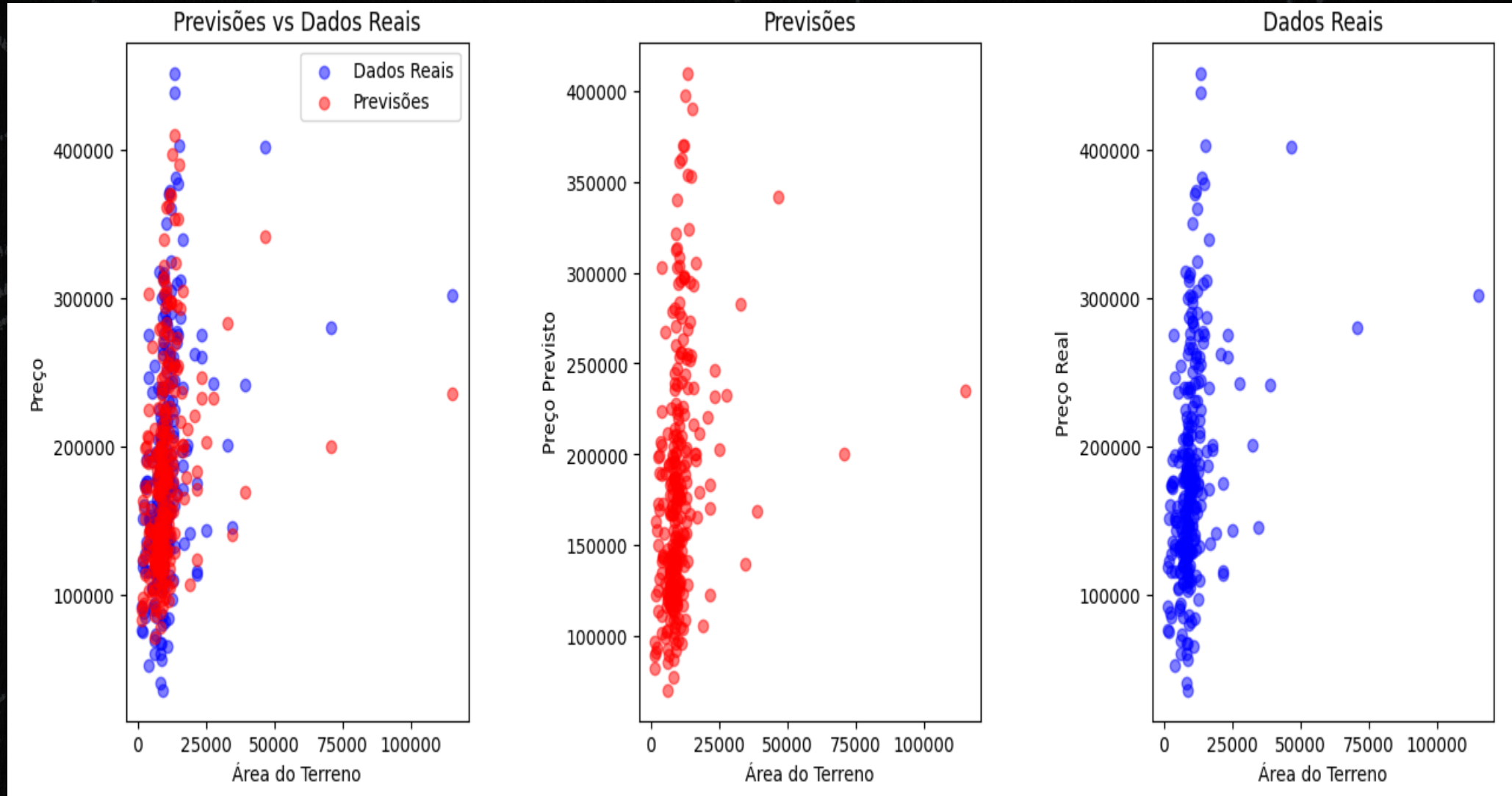
0.89

K-Fold Cross Validation:

0.83

Resultados

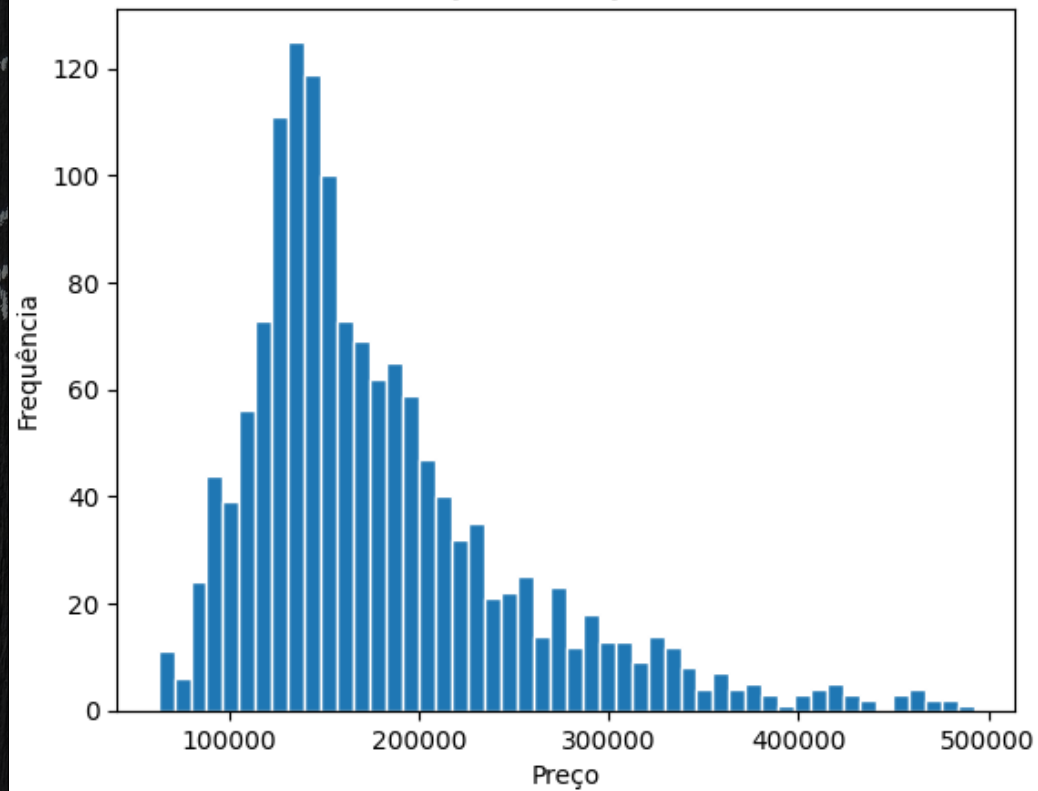
Comparação entre valores reais e previsões



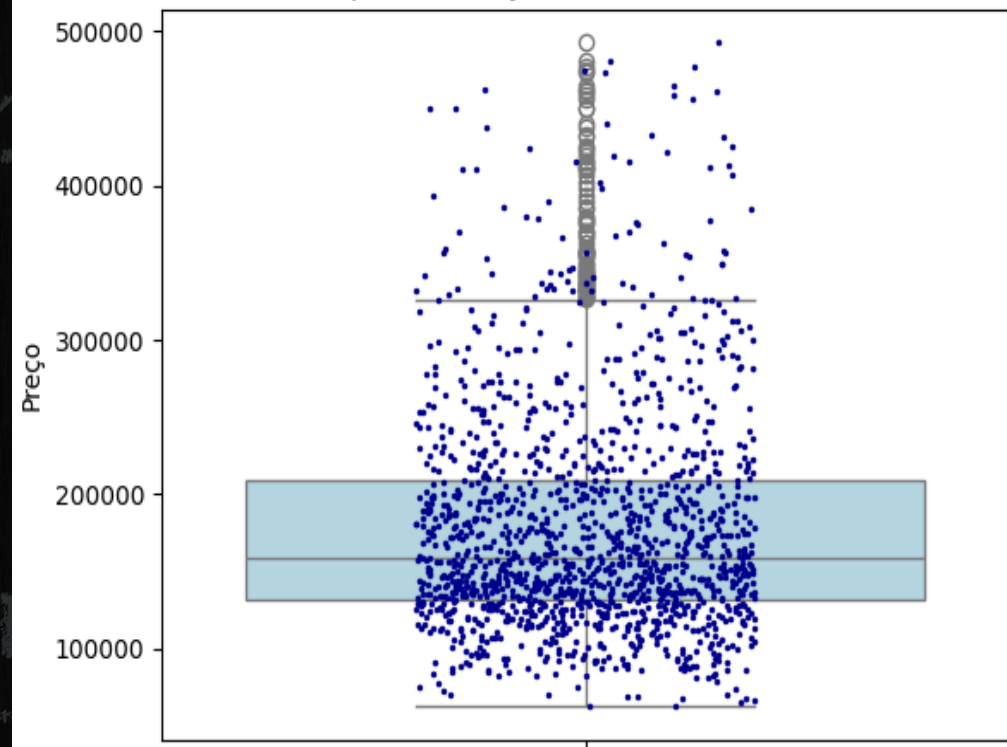
Resultados

Análises com os dados de teste

Distribuição do Preço dos Imóveis



Boxplot do Preço com Pontos Individuais



FIM

