

PUC Minas Virtual

Pós-Graduação Latu Sensu em
Ciência de Dados e Big Data

Rio de Janeiro
2021



Trabalho de Conclusão de Curso

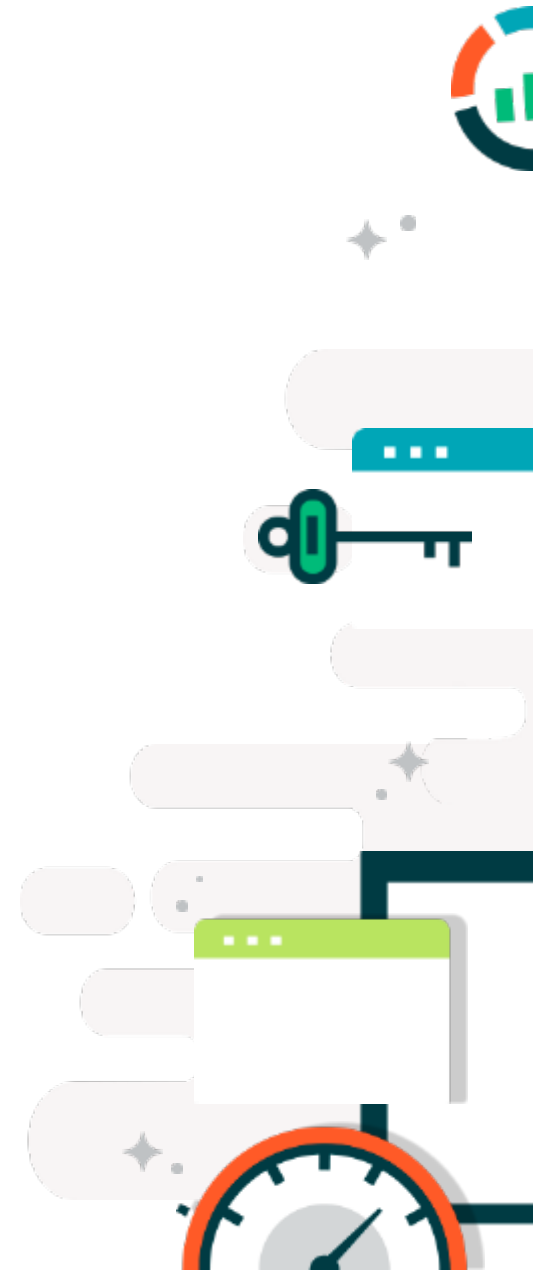
Big Deal: Encontrando hospedagens
com os melhores custos-benefícios.

Fabio Rafael Ladislei de Souza



Definição do Problema

Ordenar anúncios do AirBNB por um ranking de custo-benefício, que será calculado com base em modelos preditivos de regressão.



Coleta dos Dados

- Leitura dos dados
- Geração dos dataframes iniciais



Processamento dos Dados

- Exclusão de registros incompletos
- Agregação dos dataframes



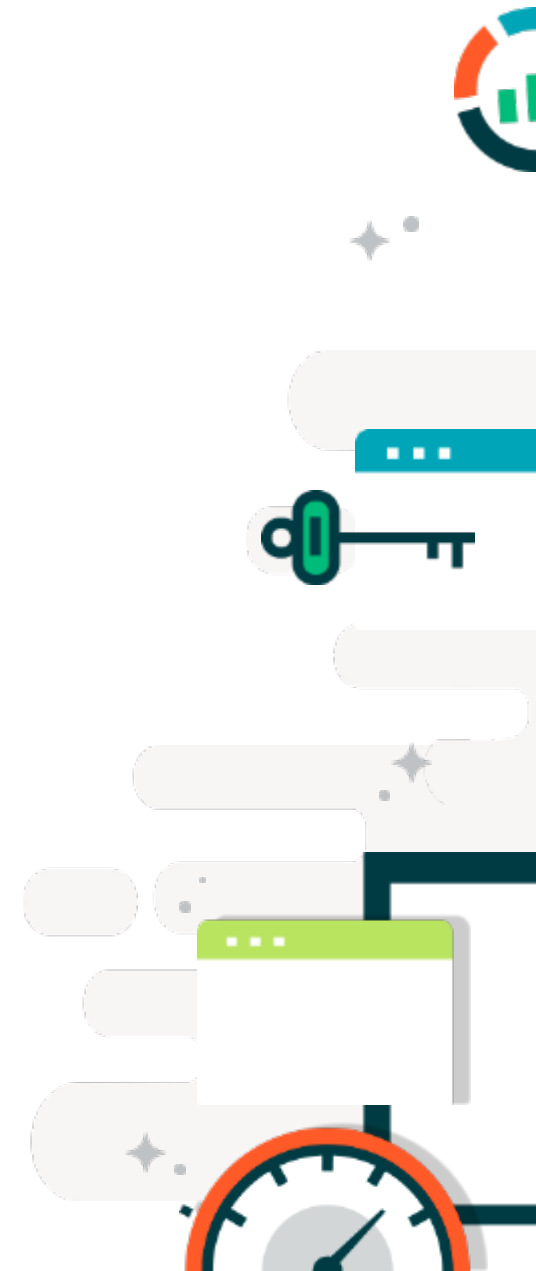
Análise Exploratória

- Exclusão de valores discrepantes
- Análise de correlação
- Seleção de atributos
- Plotagens diversas
- Limpeza dos dados
- Adequação dos dados ao aprendizado de máquina
- Matriz de correlação dos atributos



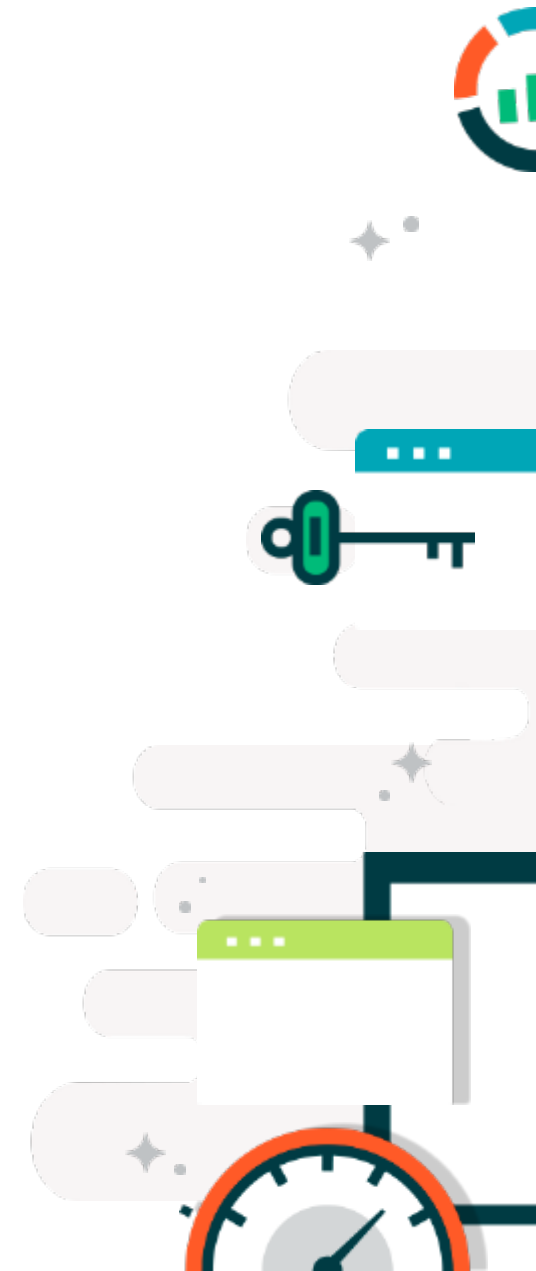
Modelagem Preditiva

- Particionamento dos dados
- Preparação de modelos preditivos
- Definição do algoritmo de base
- Preparação da validação cruzada
- Otimização dos modelos
- Seleção, treinamento e persistência do modelo



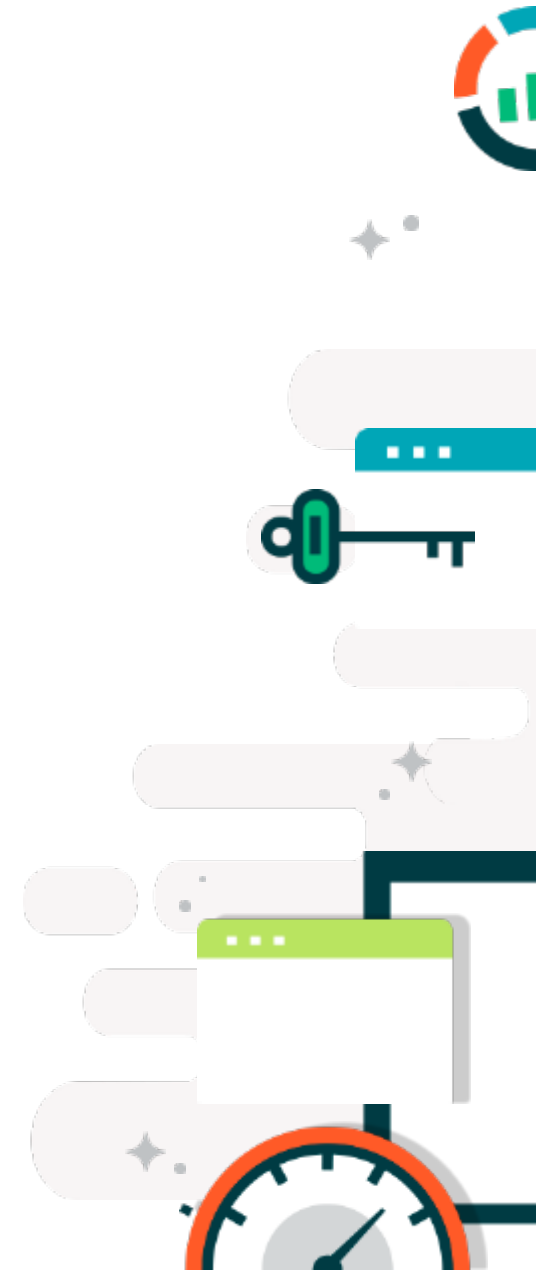
Interpretação dos Resultados

- Análise da performance do modelo
- Teste do modelo
- Visualização das previsões
- Análise dos resultados (previsto x realizado)
- Criação do ranking de custo-benefício
- Visualização dos anúncios ordenados pelo ranking



Apresentação dos Resultados

- Exibição da lista de anúncios ordenada pelo ranking
- Exibição dos dados do anúncio selecionado





Obrigado