



CASE TÉCNICO  
CIÊNCIA DE DADOS

## 1º Exercício: Classificação

<https://www.kaggle.com/c/career-con-2019/data>

CareerCon 2019 - Help Navigate Robots

Compete to get your resume in front of our  
sponsors

[www.kaggle.com](http://www.kaggle.com)

**Objetivo:** Desenvolver um modelo para prever a categoria de um conjunto de dados baseado em características fornecidas.

### Pontos Principais:

- Exploração e pré-processamento dos dados.
- Escolha de métricas de avaliação adequadas para o problema de classificação.
- Comparação de pelo menos dois algoritmos de classificação e justificativa da escolha do modelo.
- Análise de viés e variancia
- Criar endpoint, pode ser local, para predição de novas amostras.



## 2º Exercício: Regressão

<https://www.kaggle.com/competitions/competitive-data-science-predict-future-sales/overview>

Predict Future Sales

Final project for "How to win a data science competition" Coursera course

[www.kaggle.com](https://www.kaggle.com)

**Objetivo:** Criar um modelo que estime um valor contínuo baseado em várias variáveis de entrada.

### Pontos Principais:

- Análise exploratória para entender as relações entre as variáveis.
- Tratamento de valores ausentes e outliers, se necessário.
- Seleção de características e justificativa de sua importância para o modelo.
- Escolha de um modelo de regressão e explicação do motivo dessa escolha.
- Avaliação do desempenho do modelo e interpretação dos resultados.

## Combinados

Para cada exercício, por favor, documente seu processo de pensamento, a escolha das técnicas utilizadas e qualquer suposição feita durante o desenvolvimento dos modelos. Além disso, sua capacidade de comunicar suas descobertas de maneira clara e concisa será avaliada.

Por favor, envie-nos seus notebooks Jupyter com os códigos desenvolvidos e qualquer documentação adicional que você considere relevante.

Se você tiver qualquer dúvida ou precisar de esclarecimentos adicionais, não hesite em entrar em contato.

## Prazo para devolução:

O prazo para resolução é de **5 dias** a partir do recebimento do documento.



