

Objetivo

O objetivo deste projeto é aplicar os conceitos aprendidos no módulo para desenvolver um modelo de aprendizado supervisionado fazer a regressão de uma variável organizada ao longo do tempo (preço de ações, trajetória de algum sistema de direção autônomo, etc.).

Descrição

O projeto será dividido em duas partes:

- **Parte 1:** Exploração dos dados.
- **Parte 2:** Desenvolvimento e avaliação do modelo de aprendizado supervisionado.

Parte 1

Na parte 1 do projeto, vocês devem realizar as seguintes atividades:

- Explorar o conjunto de dados para entender a distribuição dos dados e identificar possíveis problemas.
- Tratar os valores faltantes;
- Separar os dados em conjunto de treinamento e conjunto de teste.

Parte 2

A parte 2 do projeto consiste no desenvolvimento e avaliação de um modelo de aprendizado supervisionado para prever o futuro de uma série temporal.

Para isso, é necessário avaliar o desempenho dos seguintes algoritmos:

- Regressão Linear
- Árvore de Decisão
- RandomForest
- XGBoost
- MLP
- LSTM

Por fim, é necessário comparar os resultados dos diferentes algoritmos para escolher o melhor modelo para o problema em questão.

Obs.: É necessário que resolva o problema criando pipelines simples e otimize os hiperparâmetros do modelo escolhido como o melhor.

Métricas de avaliação

Para avaliar os melhores modelos, sugerimos que sejam usadas as seguintes métricas:

- R^2
- RMSE
- MAE

Requisitos

Para realizar este projeto, os seguintes conhecimentos são necessitados:

- Fundamentos de aprendizado supervisionado;
- Métricas de avaliação;
- Algoritmos de regressão;
- Utilização do scikit-learn e keras

Entrega

Será necessário a entrega de um relatório do projeto (.ipynb), contendo as seguintes informações:

- Uma descrição dos dados e do problema de Regressão.
- Uma análise dos resultados dos diferentes algoritmos.
- Uma conclusão, com recomendações para a melhoria do modelo.

Avaliação

O relatório do projeto será avaliado com base nos seguintes critérios:

- Clareza da explicação.
- Precisão da análise.
- Adequação das recomendações.

Cronograma de entrega

As entregas serão semanais e devem seguir o cronograma apresentado abaixo:

Tarefa	Data de entrega
Exploração dos dados	01/08/2024
Desenvolvimento e avaliação dos modelos	09/08/2024