

Projeto Final
Mestrado Ciências de Dados para Empresas
Aprendizagem Supervisionada

(Fev 2024)

Prof. Helena Penalva & Prof. Vala Ali Rohani

Objetivo

Aplicar os métodos de Aprendizagem supervisionada na construção de modelos de previsão

Tópicos do Projeto e Conjuntos de Dados:

1. *Bike Sharing Dataset* (Regressão Linear Múltipla)

Prever o número de alugueis de bicicleta por hora, ou por dia, com base nas condições ambientais e sazonais. Considerar a variável dependente – “cnt” (excluir da análise as variáveis “casual” e “registered”).

<https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>

2. *Online Shoppers Purchasing Intention Dataset* (Classificação)

Prever se as intenções de compradores online, numa sessão, resultam numa compra. Considerar a variável dependente “Revenue”.

<https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset>

A. Objetivos do Projecto:

- Limpar o conjunto de dados.
- Realizar uma análise exploratória aos dados.
- Desenvolver e validar um modelo (Regressão ou Classificação).
- Ajustar outros modelos, melhorando a precisão face ao modelo inicial.

B. Entregas:

- Relatório do projeto em pdf.
- Código do Projeto (R).

[Por favor, compactar todos os ficheiros (ZIP) e enviar por e-mail helena.penalva@esce.ips.pt]

C. Pontuação do Projecto:

O projeto será avaliado de acordo com o seguinte:

Introdução [2,0 Valores]

[1,0 Valor] Definição dos objetivos.

[1,0 Valor] Apresentação do conjunto de dados e identificação das variáveis dependente e independentes.

Limpeza de Dados e Análise Exploratória de Dados [5,0 Valores]

[5,0 Valores] Carregar os dados no R e limpá-los; armazenar os dados limpos num ficheiro separado categorizar as variáveis categóricas; Análise exploratória de dados.

Métodos de Aprendizagem Supervisionada (Regressão Múltipla ou Classificação) [8,0 Valores]

[1,0 Valor] Preparação dos conjuntos de dados (dados de Treino e de Teste).

[4,0 Valores] Desenvolvimento dos modelos.

[3,0 Pontos] Comparação dos modelos desenvolvidos.

Conclusões [3,0 Valores]

[3,0 Valores] Discussão dos resultados e conclusões.

Qualidade do relatório do projeto [2,0 Valores].

D. Datas importantes

Data de Entrega

17 de março de 2024
