



Relatório de Análise de Modelos de Regressão

Universidade Federal do Espírito Santo
Probabilidade e Estatística

Thales Melo da Penha

6 de Novembro de 2023

Objetivo

Este relatório descreve a análise de modelos de regressão para prever a variável Y com base nas características X1, X2, X3 e X4. O código que eu fiz em Python carrega os dados do arquivo "Trabalho.csv", divide-os em treino(70%)/teste(30%), ajusta quatro modelos (KNN com K=1, 3, 5 e Regressão Linear), calcula o MAE para cada modelo e estima intervalos de confiança (95%) para o MAE.

Resultados

A figura a seguir apresenta a análise dos resultados:

```
thales@Thales-VM: ~/Área de trabalho/PROB_EST/trabalho_probabilidade
(myenv) thales@Thales-VM:~/Área de trabalho/PROB_EST/trabalho_probabilidade$ python ProbEst-Trabalho.py
----- Erro Absoluto Médio (MAE) -----
Modelo 1 - KNN (K=1) --> [0.66]
Modelo 2 - KNN (K=3) --> [0.67]
Modelo 3 - KNN (K=5) --> [0.70]
Modelo 4 - Regressão_Linear --> [0.45]

----- Intervalos de Confiança (95%) -----
Modelo 1 - KNN (K=1) --> (0.43, 0.88)
Modelo 2 - KNN (K=3) --> (0.46, 0.88)
Modelo 3 - KNN (K=5) --> (0.46, 0.93)
Modelo 4 - Regressão_Linear --> (0.36, 0.54)

(myenv) thales@Thales-VM:~/Área de trabalho/PROB_EST/trabalho_probabilidade$
```

Figura 1: Resultados gerados pelo código Python

Conclusões

- O modelo de Regressão Linear obteve o menor Erro Absoluto Médio (MAE) em comparação com os modelos KNN com diferentes valores de K. Isso indica que a Regressão Linear apresenta um melhor desempenho na previsão da variável Y com base nas características X1, X2, X3 e X4.
- Os modelos KNN com K = 1, 3 e 5 mostraram MAEs mais altos, sugerindo que eles têm uma precisão ligeiramente menor na previsão em comparação com a Regressão Linear.
- Os Intervalos de Confiança (95%) indicam a variabilidade nas estimativas do MAE. A Regressão Linear apresentou um intervalo de confiança mais estreito, o que sugere maior confiança na estimativa do MAE, enquanto os modelos KNN tiveram intervalos de confiança ligeiramente mais amplos.
- A Regressão Linear é mais simples e interpretável, enquanto os modelos KNN podem ser úteis em casos de maior complexidade.

Discussões

Neste estudo, a Regressão Linear se destacou com um MAE mais baixo comparado ao KNN, mostrando maior precisão. No entanto, é importante considerar fatores específicos para escolher o melhor modelo, uma vez que a diferença entre eles não foram suficientemente grandes.