

PAD Individual Report

João Lopes - 60055

No âmbito da disciplina de Prospeção e Análise de Dados, eu, em colaboração com o meu colega José Romano, concluímos de forma bem sucedida o projeto proposto.

Desde o início, delineamos uma estratégia de trabalho em que eu assumi a responsabilidade pelo primeiro ponto, centrado na regressão linear, enquanto trabalhamos em conjunto no terceiro ponto, relacionado com o Fuzzy Clustering. O meu colega ficou encarregado da análise da implementação de PCA.

A minha experiência em regressão linear, adquirida através da especialização em Inteligência Artificial, foi fundamental para o sucesso do módulo correspondente. Iniciamos o processo com uma análise detalhada do conjunto de dados selecionado, identificando possíveis variáveis que poderiam influenciar os resultados finais. Destacamos especialmente a atribuição de um índice de cliente a cada utilizador de cartão de crédito, reconhecendo o impacto potencial disso nas análises subsequentes.

Além disso, realizamos uma investigação metódica da integridade dos dados, identificando e tratando valores nulos através de métodos de inferência, garantindo assim a robustez do conjunto de dados utilizado.

Dado que ao contrário do meu colega eu não tenho bases suficientemente fortificadas de finanças/economia, optei por analisar a correlação entre as variáveis para selecionar as mais relevantes para a regressão linear. Este processo envolveu o uso inicial de bibliotecas como sci-kit learn em Python, porém, optamos por uma abordagem mais detalhada para garantir uma compreensão aprofundada dos cálculos envolvidos.

Concluindo o tópico da regressão linear posso afirmar que não só consolidei o meu entendimento dos conceitos fundamentais de análise de dados, mas também a minha capacidade de aplicar esses conceitos na prática. Estou confiante de que as conclusões obtidas por este projeto contribuirão significativamente para o nosso conhecimento e para a área de estudo em questão.

Relativamente à última parte, deparámo-nos com uma série de problemas e contratempos que culminou numa resolução incompleta do FCM, nomeadamente relativamente ao IAP, o que nos levou a tomar uma abordagem mais cautelosa e improvisada de modo a manter a priorizar a coerência das ideias apresentadas ao longo do projeto, minimizando ao máximo os percalços que tivemos de ultrapassar.

Em essência, depois de ter concluído o primeiro trabalho referente à cadeira de Prospeção e análise de dados posso concluir que o mais importante é mesmo perceber com que intuito são utilizadas as contas efetuadas de forma a registar resultados e como é que estas são elaboradas, assim como entender que não existem dois datasets iguais e que as

conclusões retiradas destes são sempre únicas e que nem sempre os datasets se encontram preparados para serem estudados, sendo necessário fazer o pré-processamento dos dados.