

Descrição do problema

Dentro da empresa existe uma equipe que é responsável por definir as características de produtos que vão ser construídos. Para uma categoria de liquidificadores, por exemplo, essa equipe fica responsável por definir a cor, volume do copo, material do copo ou mesmo se vai ter um filtro para suco de frutas.

Para definir as características das diferentes categorias de produtos, a equipe procura entender quando o cliente se interessa por um produto. É nesse ponto que o time do Labs é chamado. O seu objetivo principal é desenvolver um modelo que consiga classificar o produto entre "interessante" e "não interessante" dados os atributos do mesmo. Além disso, é desejável que você também consiga indicar quais são os atributos e respectivos valores que mais impulsionam o interesse do cliente.

Mais importante que o resultado final é a sua estratégia de solução. Deixe claro cada passo da sua solução e explique as decisões tomadas durante o processo. Também é importante que você nos mande o seu código desenvolvido (Python ou Scala). Uma sugestão é usar Jupyter Notebook. Assim, você consegue explicar cada ponto da sua estratégia de forma mais explícita.

Esperamos receber de você um documento com a explicação do trabalho desenvolvido.

Bom trabalho!

Conjunto de dados

O conjunto de dados possui os atributos de cada produto de uma categoria de painéis e se houve ou não interesse do cliente ao visualizar o produto no site. Cada linha é o registro da visualização de um cliente em determinado momento.