

Desafio para a vaga de IA Engineer

Bemol Digital

Olá, parabéns por ter chegado a essa nova fase. Nós elaboramos um desafio bem bacana para conhecermos melhor a sua forma de trabalho.

No varejo, é muito importante conhecer o perfil do nosso público. Uma segmentação de clientes bem definida pode ajudar na alocação eficaz de recursos de *marketing*, favorecendo a construção de um relacionamento saudável e de longo prazo com os clientes.

Com o objetivo de fornecer um melhor atendimento aos nossos consumidores, gostaríamos de projetar algumas estratégias de vendas e marketing e, para isso, contamos com a sua ajuda. Logo abaixo temos alguns links para obter conjuntos de dados com transações de venda em lojas de varejo. Lembre-se que são apenas sugestões, se encontrar outro que se interesse mais, pode ficar à vontade.

- Kaggle: “Retail Transaction Data” e “Online Retail K-means & Hierarchical Clustering”.

Seu trabalho aqui vai ser segmentar uma lista clientes baseados em suas transações. Não esqueça de nos contar sobre o método escolhido, as métricas utilizadas e exibir a visualização dos grupos que foram formados, nos mostrando suas principais características.

Em seguida, essas informações vão ser utilizadas pela equipe de vendas e *marketing*. Devido a algumas mudanças na política de segurança da empresa, no que diz respeito a dados, agora só é permitido consumir dados via API. Partindo do ponto em que você já passou pelo processo de conseguir os dados, fez a segmentação e tem seus resultados em mãos, **sua missão é disponibilizar esses dados para o restante do time via API.** A entrada será um ou vários *ids* de clientes e a saída será o *id* e *segmento*.

Não temos restrição quanto a *stack*, mas gostamos do básico, algo como: Python, SQL, Node.js. Sinta-se livre para utilizar sua *stack* preferida.

Alguns diferenciais no seu projeto:

- Mostrar os artefatos gerados durante o processo de construção (códigos, arquitetura, programas utilizados)
- API documentada
- Projeto versionado usando Git (para API)
- Metadados da base de dados (pode ser um .docx, planilha online, etc)
- Disponibilizar essa aplicação em alguma nuvem: Azure, Google Cloud, AWS (eles possuem *trial*)
- Mostrar um exemplo de consumo da API, poderia ser no jupyter(lab/notebook)

Dica: Menos é mais, *keep it simple* :)