



Desafio para a vaga de AI Engineer

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 vendas 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 consiste em segmentar uma lista de clientes baseada em suas transações. Em seguida, essas informações serão 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 seu respectivo *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. Como ocorre na maioria dos problemas em *data science*, algumas explicações sobre a solução podem ser requeridas durante o processo de seleção. Não esqueça de nos contar sobre o *dataset* escolhido, o método de segmentação, as métricas utilizadas, exibir visualizações acerca dos grupos formados e demais análises, nos mostrando suas principais características. É importante que você deixe o mais claro possível suas suposições e decisões de projeto feitas ao longo do desenvolvimento do desafio.

O que não pode faltar no seu projeto:

- Relatório detalhado com análises exploratórias, métodos e métricas utilizados, além de decisões de projeto tomadas ao longo do desenvolvimento do desafio - pode ser feito no próprio jupyter(lab/notebook);
- Mostrar os artefatos gerados durante o processo de construção (códigos, arquitetura, programas utilizados, etc.);
- Metadados da base de dados (pode ser um .docx, planilha online, etc.);
- API documentada;
- Fornecer todo o direcionamento para execução e mostrar um exemplo de consumo com a API.

O que seria um diferencial no seu projeto:

- Projeto versionado usando Git;
- Disponibilizar essa aplicação em alguma nuvem: Azure, Google Cloud, AWS (eles possuem *trial*);
- Testes unitários;
- Empacotamento da aplicação.

Tenha um ótimo desafio e divirta-se! :)

Time de Inteligência Artificial