



Data Science Academy

www.datascienceacademy.com.br

Formação Cientista de Dados

Projeto com Feedback 10

Sistema de Recomendação Para Rede de
Varejo Usando Market Basket Analysis

Quer você faça compras com listas de compras meticulosamente planejadas ou deixe que o capricho guie seus passos, nossos rituais únicos de compra definem quem somos. Instacart, um aplicativo de pedido e entrega de supermercado, tem como objetivo facilitar o preenchimento de sua geladeira e despensa com seus itens pessoais favoritos e itens básicos quando você precisar deles. Depois de selecionar produtos por meio do aplicativo Instacart, os compradores revisam seus pedidos, fazem compras e a entrega é feita na loja mais próxima a você.

A equipe de ciência de dados da Instacart desempenha um papel importante no fornecimento dessa experiência de compra agradável. Atualmente, eles usam dados transacionais para desenvolver modelos que preveem quais produtos um usuário comprará novamente, quais tentará pela primeira vez ou quais adicionará ao carrinho durante uma sessão. Recentemente, a Instacart disponibilizou esses dados de forma aberta e o link para download você encontra logo abaixo

Neste projeto de ciência de dados, você usará esses dados anônimos nos pedidos dos clientes ao longo do tempo para prever quais produtos adquiridos anteriormente estarão no próximo pedido de um usuário.

Nossa recomendação é que você utilize linguagem R e os pacotes de Market Basket Analysis oferecidos em R. O link para download do dataset você encontra aqui:

<https://www.instacart.com/datasets/grocery-shopping-2017>

Aqui você encontra um post do VP da Instacart definindo o problema em detalhes:

<https://tech.instacart.com/3-million-instacart-orders-open-sourced-d40d29ead6f2>

Quando concluir o projeto, envie os scripts e datasets para projeto@dsacademy.com.br. Caso os datasets usados sejam muito grandes, armazene em um diretório virtual (existem vários na internet, como Google Drive ou Dropbox) e envie o link para que nossa equipe possa baixar os datasets. Se os arquivos foram pequenos (uma amostra do dataset original), envie no anexo junto com o script. Documente seu script tanto quanto possível.

Caso prefira, disponibilize seu projeto no Github e envie o link do seu repositório para nossa equipe no e-mail projeto@dsacademy.com.br. Nesse caso, o Readme do repositório deve constar que este trata-se de um projeto da Formação Cientista de Dados da Data Science Academy.

Em até 24 horas, daremos o feedback!

Bom trabalho!