

ABI AcademyHack Ambev



Julia Gontijo Lopes
Marcus Navarro Gabrich
Rodrigo de Oliveira Gomes

SISTEMA DE RECOMENDAÇÃO DE DESCONTOS PERSONALIZADOS



O PROBLEMA:

- Para quem conceder desconto?
- Como saber se o desconto foi aproveitado?
- Como saber qual desconto terá retorno positivo?

A SOLUÇÃO:

- Personalizar descontos
- Medir seu retorno
- Personalizar novamente

OBJETIVO

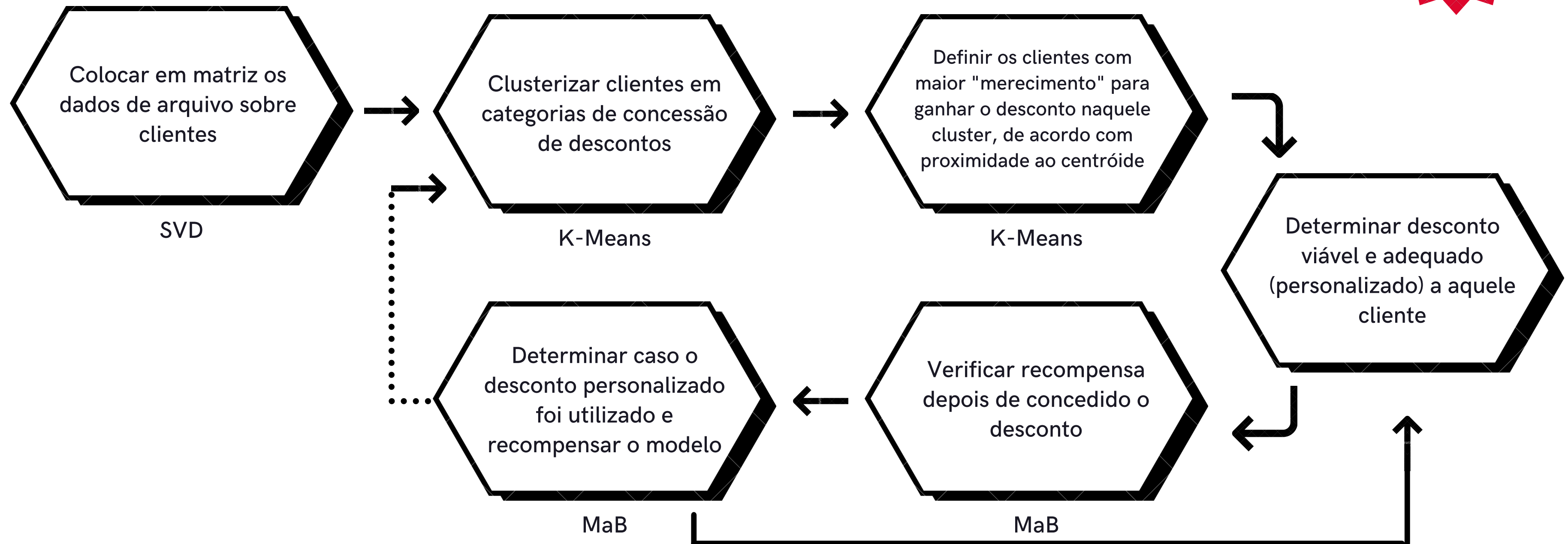
Fornecer descontos que revertam em compras futuras
- e potencialize a possibilidade de uma compra ser feita -

quantidade, produto, volume

Desenvolver sistema que cria um modelo do perfil de usuário e a partir dos dados, gera suposições desse padrão de compra, e define, segundo esses padrões, qual evento de desconto trará maior recompensa á Ambev e qual será mais atrativa ao cliente.

compras futuras, aumento de quantidade de produtos
comprados, adoção á compra de novos produtos

DESENHO DA ARQUITETURA DA SOLUÇÃO DESENVOLVIDA



LISTA DE FERRAMENTAS INTEGRADAS:



(import os, import sys, import csv, import json, import io)
tratamento de data set de entrada.

(import pandas)
leitura do dados

(import matplotlib.pyplot, import seaborn)
para demonstrar dados graficamente

(import numpy)
tratamento de dados numéricos

(import utils)
ferramentas gerais para auxilio no funcionamento das
diversas técnicas embutidas

(from sklearn.model_selection import
ShuffleSplit
from sklearn.cluster import KMeans
from sklearn.metrics import
silhouette_samples, silhouette_score)

Bibliotecas necessária para aplicar técnicas de
processamento de sinal afim de tratar os
dados

(import pickle) salvar resultado das técnicas
para usos futuros

BIBLIOTECAS ALTERADAS UTILIZADAS:



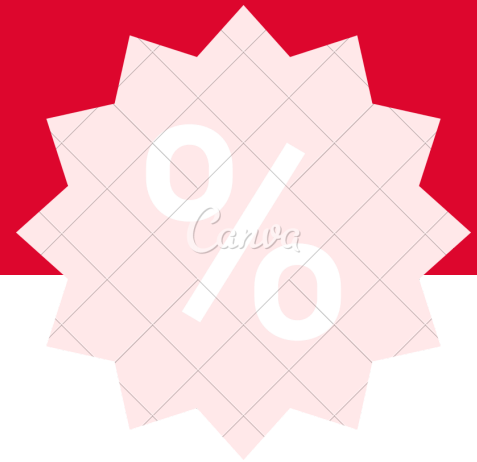
(from funk_svd.svd)

versão alterada do svd(Singular value decomposition) afim de facilitar manuseio

(import mab)

versão alterada do MaB (multi-armed bandit) afim de gerar resultados de valor em relação ao numero de acertos efetuados

BASES DE DADOS INICIAIS:



- "Original_data.csv": Base de dados original disponibilizada
- "DataTreated.csv": Base de dados alterada segundo base disponibilizada pela Ambev com os atributos:

Ship-to ID	Volume_2019	Volume_2018	sfdc_tier	segment	Buying_Average	Became_Member_on	Volume_2019_MNP	Most_Wanted_Product(MNP)
---------------	-------------	-------------	-----------	---------	----------------	------------------	-----------------	--------------------------

Concedidos pela Ambev

Média de compras em relação á visitas no site de um cliente. Dado utilizado para definir se um cliente é um "carrinho abandonado" em potencial caso sua relação de compras/visitas esteja abaixo da sua média normal

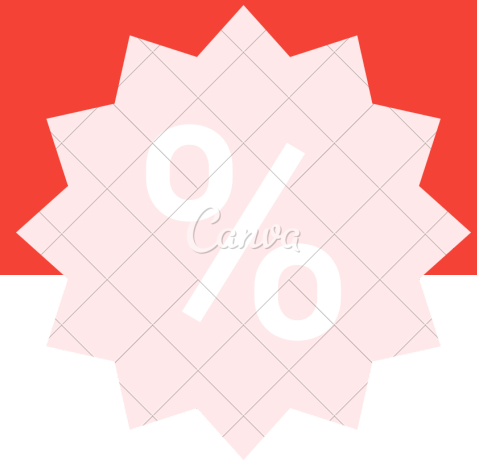
Concedidos pela Ambev

data em que a POC se tornou um cliente para maior definição de fidelidade

o produto mais comprado por um cliente (em relação ao volume de compra de cada um em 2019), tido como produto preferido do cliente.

*dados fictícios

FUNCIONALIDADE



O algoritmo desenvolvido seria uma integração á plataforma de e-commerce da Ambev "BEES" para definir qual desconto disponibilizar aos clientes cadastrados através da análise de qual desconto gera maior retorno (lucro/vendas futuras), ou seja, recompensa, á própria Ambev caso seja disponibilizado a aquele cliente específico.

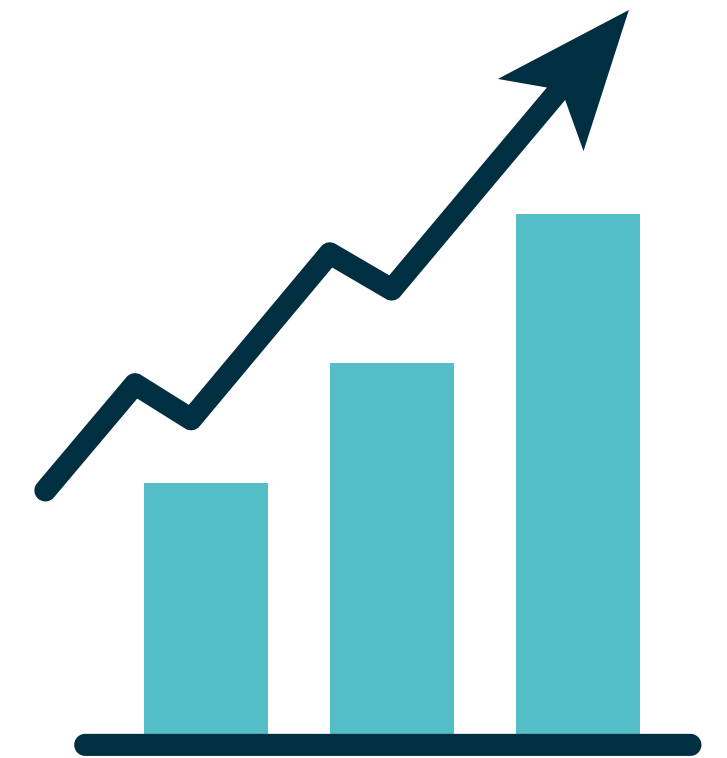
FUNCIONAMENTO DO SISTEMA:

- Um cliente irá realizar uma compra na plataforma "BEES"
- Ao selecionar uma compra, o sistema calcula qual desconto provavelmente o cliente irá usar e irá trazer recompensa á Ambev
- Este calculo é feito pela análise de seu perfil de compra de acordo com seu histórico na plataforma.



FUNCIONAMENTO DO SISTEMA:

- O desconto é concedido
- O sistema analisa se o cliente viu o desconto, se usou o desconto, se finalizou a compra, e, em sua próxima vinda á plataforma, computa se este realizou mais compras (compras futuras)
- Com essas informações, o sistema consegue analisar se o desconto foi o ideal para aquele cliente, e ajustar qual seria o ideal em uma próxima concessão de desconto.



MODELO IMPLEMENTADO:

Utilização de: SVD; KMeans e Multi-armed bandits

Manipulação e processamento dos dados

- Criação de dataset com o perfil de cada cliente e características de compra deste, para poder clusteriza-los
- Conversão dos dados não numéricos em numéricos para poder utilizar o algoritmo KMeans

SVD - transformação dos dados em uma matriz

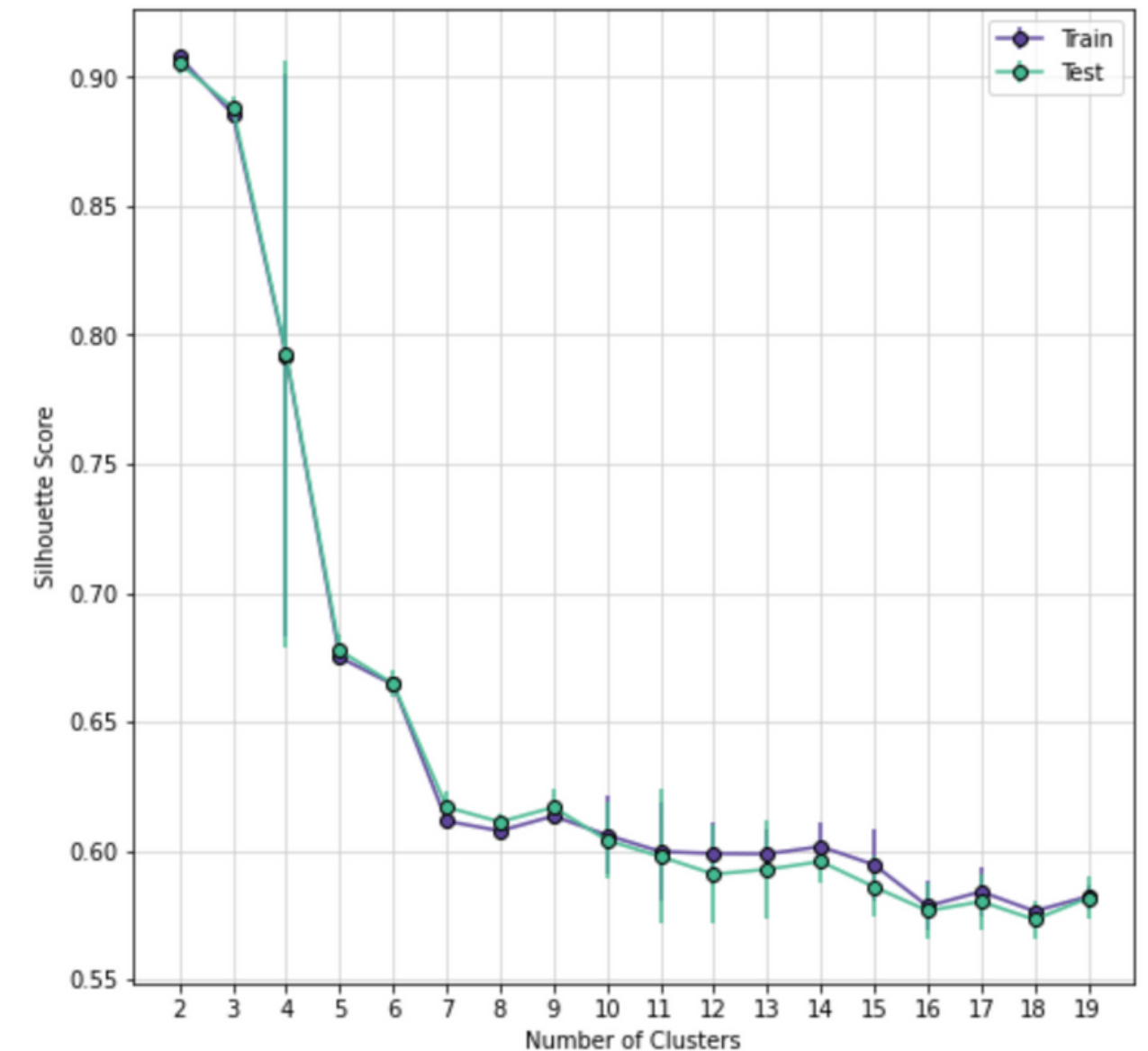
- .n-sweep: roda 5 vezes p analisar o numero apropriado de fatores para serem utilizados, baseado no rmse,my,precisionk e recallk.
- .k-sweep: com o numero de fatores estabelecidos para o LEVARGAGE, consegue analisar o quao bem o SVD da matriz terá de performance dentro do algoritmo para diferentes rankings de recomendação

MODELO IMPLEMENTADO:

Utilização de: SVD; KMeans e Multi-armed bandits

KMeans (clusterização dos clientes)

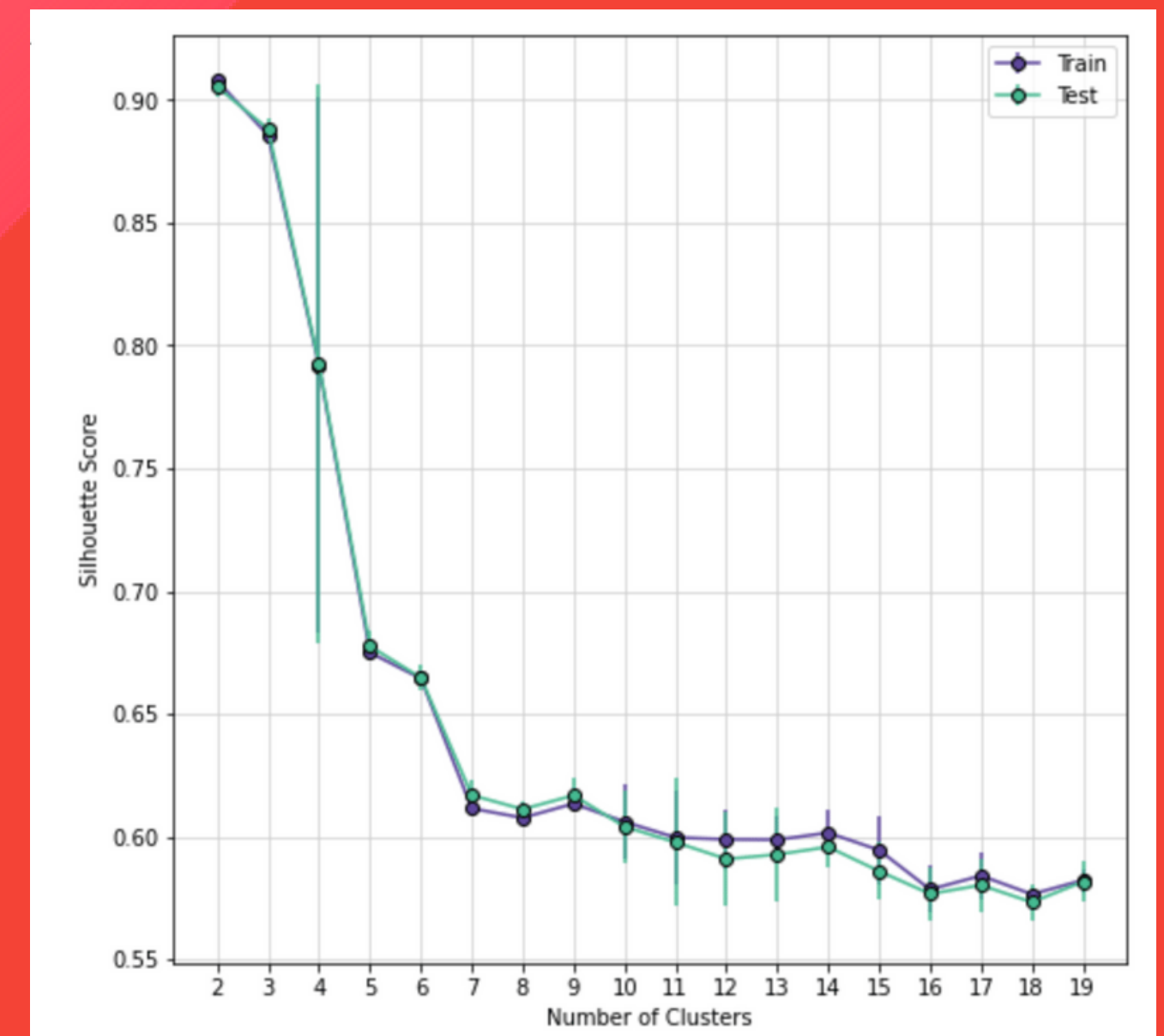
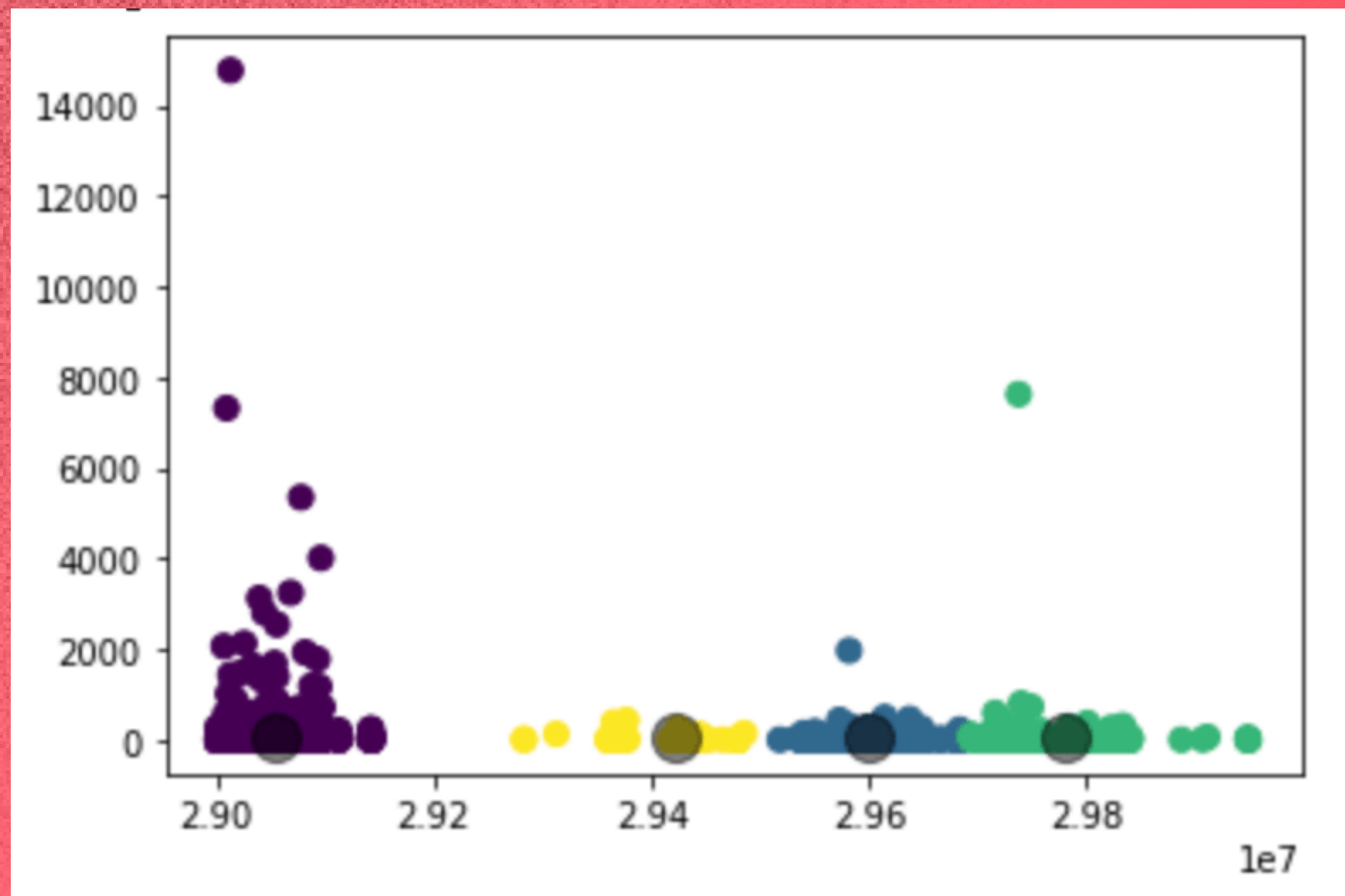
- Análise do número de clusters ideal a ser utilizado pelos "Silhouette Average"
- Pelo gráfico, podemos ver que o número ideal seria 4
- Roda o KMeans com número de clusters definido (4) e treina o modelo, definindo os agrupamentos de clientes com características e padrões de compras comuns



MODELO IMPLEMENTADO:

Utilização de: SVD; KMeans e Multi-armed bandits

Clusters formados



MODELO IMPLEMENTADO:

Utilização de: SVD; KMeans e Multi-armed bandits

MaB

- Dado um cliente X, o MaB inicia com uma recomendação aleatória. Na medida em que o sistema ganha experiência (recompensas), ele cria um padrão de descontos para cada tipo de cliente específico.

Em uma próxima compra

- Dado o desconto, o sistema consegue dar uma pontuação á aquele desconto para aquele cliente e ajustar/recalcular qual desconto seria melhor para ele e qual traria mais recompensa á Ambev pelas métricas: se o cliente viu o desconto, se usou o desconto, se finalizou a compra, e se houve compras futuras

VALOR

O cálculo da pontuação de se o desconto foi ideal para o cliente é feita da seguinte maneira:

$$\textbf{Reward} = \textbf{discount_viewed} + \textbf{discount_used} * \textbf{purchase_completed} + 2 * \textbf{Future_Purchase}$$

O peso colocado em "Future_Purchase" é maior, o que faz contar muito se o uso de desconto levou ao cliente realizar uma compra futura.

Assim, descontos que levam os clientes a comprar mais são mais pontuados, o que os leva a serem mais escolhidos no ajuste que o sistema faz e traz á Ambev maior lucro monetário.

PROPOSTAS DE NOVAS FUNCIONALIDADES E MELHORAS FUTURAS

Podemos propor uma automatização e autonomia do sistema que se divide em duas partes



Sistema utiliza de informações antigas

Ja define o usuário em um cluster específico previamente gerado e em seguida gera um desconto de entrada

A cada adição de x clientes reestrutura o sistema

Através de um modelo matemático definir um numero x de novos cliente que quando alcançados reestrutura os dados, refazendo o processo de cluterização para definir novas classificações de clientes e manter um alto numero de sucesso em vendas.

PROPOSTAS DE NOVAS FUNCIONALIDADES E MELHORAS FUTURAS

Quando conceder o desconto?

O sistema é capaz de decidir e ajustar qual seria o melhor desconto a conceder ao cliente. Futuramente, poderia ser adicionada um cálculo de quando conceder o desconto e quando não conceder



LINK DO GITHUB

<https://github.com/marcusng8/Discount-Leverage>