

Introdução à Ciência dos Dados

Aula 04

Gabriel Pacheco

Bacharel em Sistemas de Informação (UFMG)

Mestrando em Ciência da Computação (UFMG)

Trabalha em Digicade desde 2016





Curiosidade

Novo ministro da saúde (Teich) sobre teste em massa para Covid-19:

"O que eu tô te mostrando é o seguinte: mais do que qualquer coisa, a sabedoria de ter o dado e interpretar o dado, e tomar ações a partir disso, vai fazer toda a diferença"

"... O que importa é como é que você conduz de acordo com o que você tem a partir do teste."

Fonte: <https://g1.globo.com/bemestar/coronavirus/noticia/2020/04/22/nao-tem-teste-em-massa-diz-ministro-da-saude-sobre-como-sera-avaliado-o-avanco-do-coronavirus-no-brasil.ghtml>

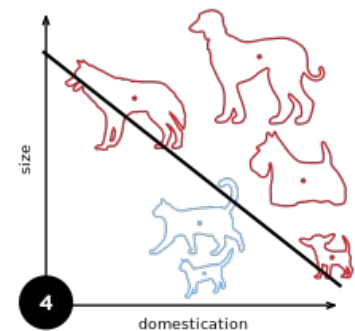
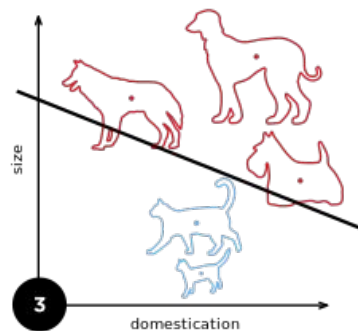
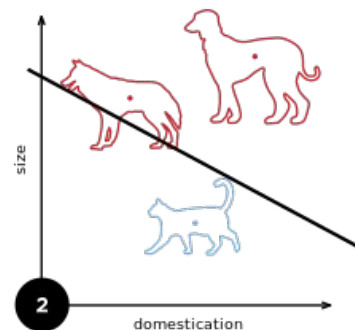
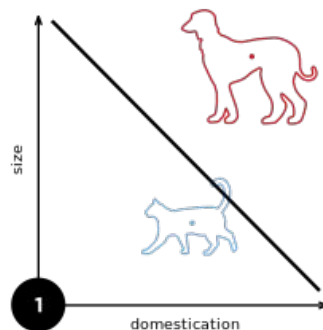


Cronograma

- Revisão (<https://www.dropbox.com/s/9k2j77h2d17zhI3/icdAula03.mp4?dl=0>)
- Associação
- Agrupamento
- Exemplo prático

Revisão Classificação

- Detecção de fraudes
- Crédito bancário
- Detecção de Spam
- Análise de Sentimentos
- Reconhecimento de imagem
- Reconhecimento de vídeo
- Análise de ações do mercado financeiro





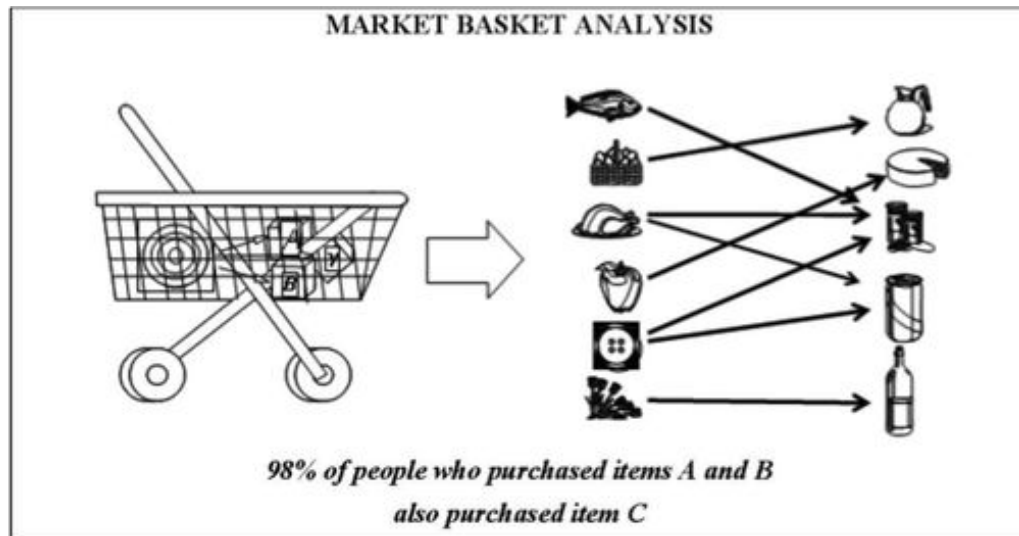
Aula 04

- Regras de associação
- Agrupamento
- Case utilizando regras de associação e agrupamento
- Exemplo prático*
- Encerramento



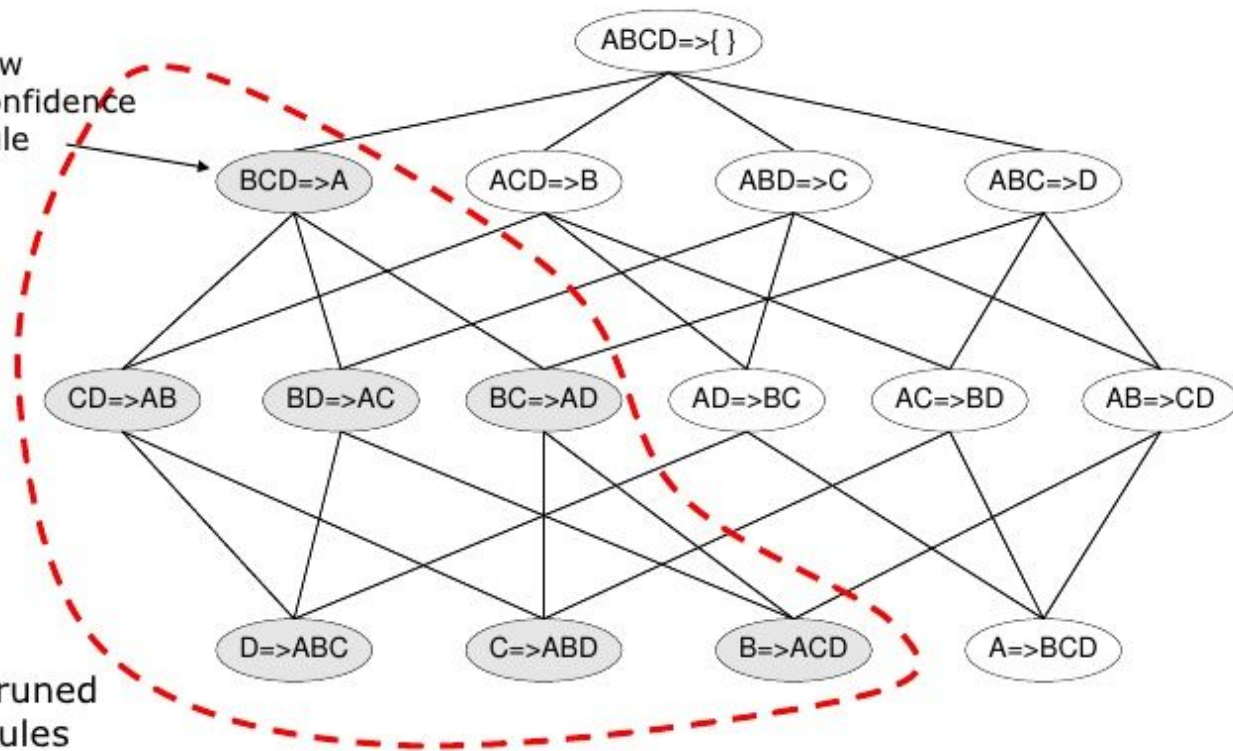
Associação

- Recomendação de produtos
 - Filmes
 - Música
 - E-commerce
- Bioinformática
- Árvores...



Low
Confidence
Rule

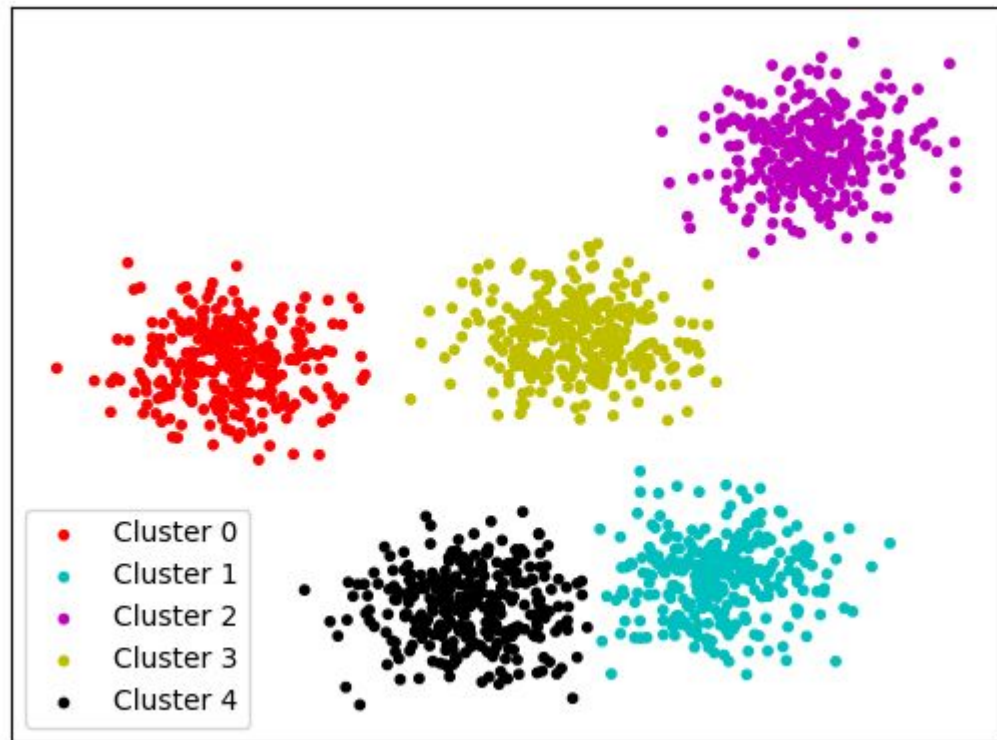
Pruned
Rules

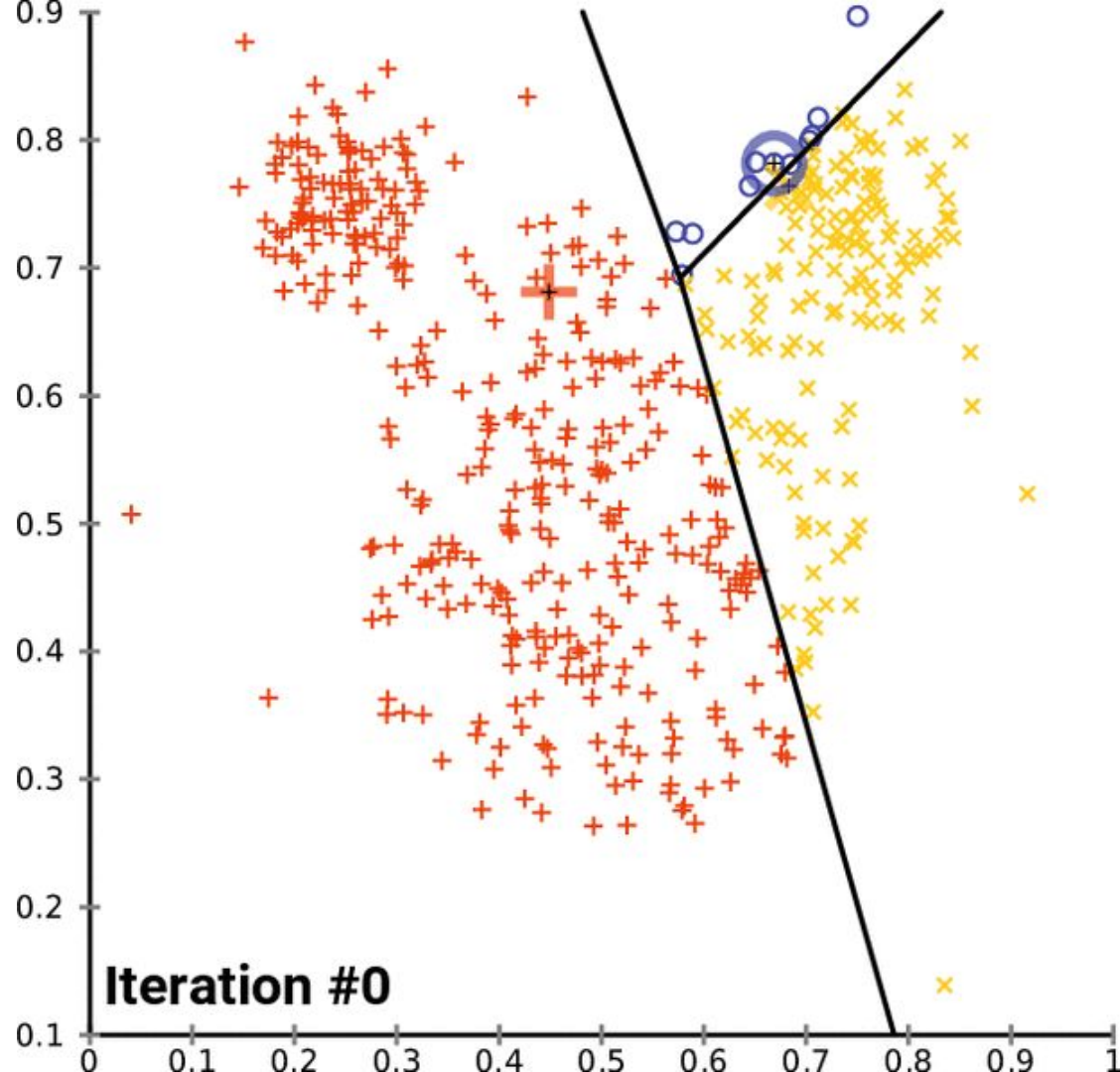




Agrupamento

- Reconhecimento de padrões
- Recuperação de informação
- Detecção de comunidades







Case: Queda de árvores em Belo Horizonte

- Artigo publicado no GeoInfo 2019 (<http://www.geoinfo.info/geo20191.html>)
- Base de dados com 300.000 árvores em Belo Horizonte
- Mais de 50 atributos
- Objetivo: predição de quedas de árvores



Exemplo prático: Iris Species

- Toy example
- Serve para aplicarmos técnicas de regressão, classificação e agrupamento
- <https://www.kaggle.com/uciml/iris>



Tutoriais e exemplos de uso de associação

- <https://www.kaggle.com/msp48731/frequent-itemsets-and-association-rules>
- <https://www.kaggle.com/datatheque/association-rules-mining-market-basket-analysis>



Tutoriais e exemplos de uso de agrupamento

- <https://www.kaggle.com/roshansharma/mall-customers-clustering-analysis>
- <https://www.kaggle.com/xvivancos/tutorial-clustering-wines-with-k-means>
- <https://www.kaggle.com/vjchoudhary7/kmeans-clustering-in-customer-segmentation>



Encerramento

- Materiais do curso: <https://github.com/gapacheco/ICD.git>



Contato

- Email: gacampacheco@gmail.com
- LinkedIn: www.linkedin.com/in/gabriel-oc-pacheco