

# Trabalhos Práticos - TP8

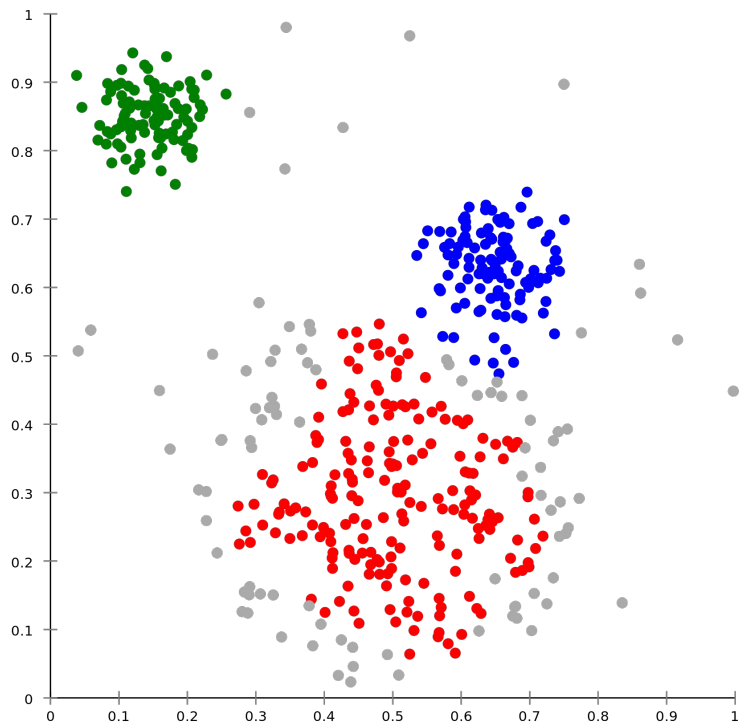
## Clusterização

Alberane Lúcio Thiago da Cunha - [alberane.cunha@posgrad.ufla.br](mailto:alberane.cunha@posgrad.ufla.br)

William Gaudência REzende - [william.rezende@posgrad.ufla.br](mailto:william.rezende@posgrad.ufla.br)

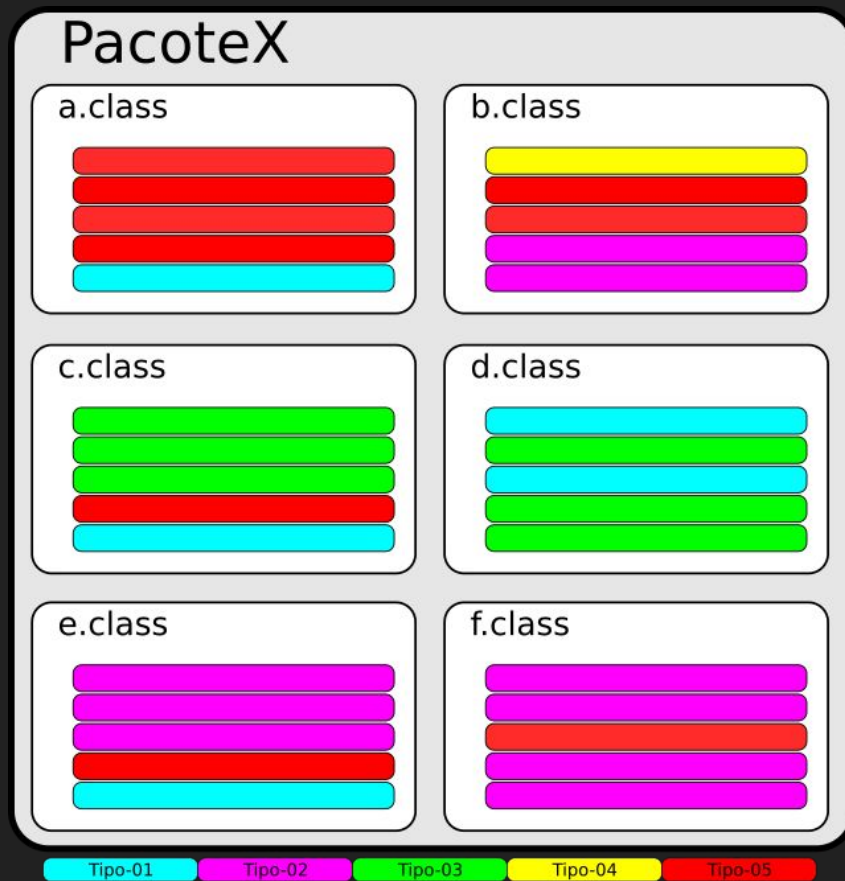
# Sobre Clusterização

É o conjunto de técnicas de **prospecção de dados** (data mining) que visa fazer **agrupamentos automáticos** de dados segundo o seu grau de semelhança.



# O Problema

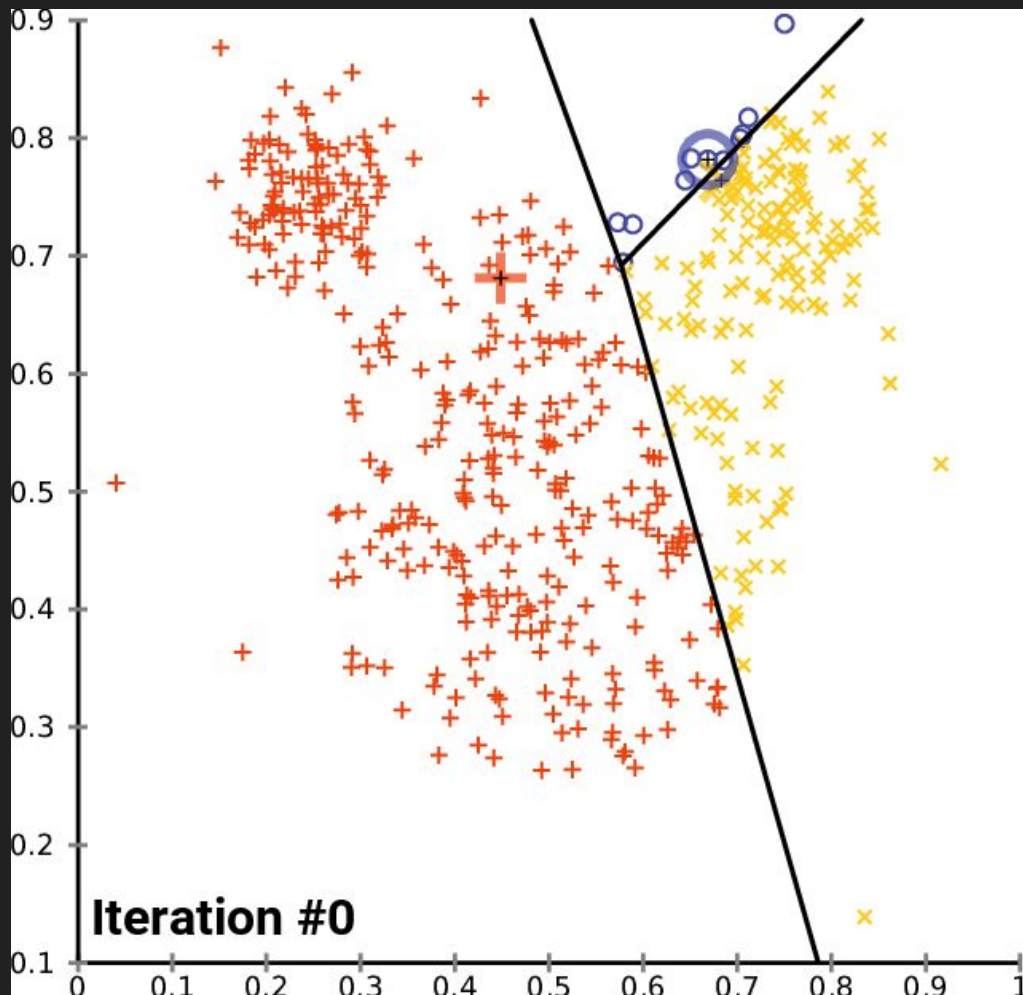
**Analisar** um conjunto de classes dentro de um único pacote e com base na análise de **todos** os tipos dessas classes propor a divisão em **N novos pacotes** baseado na **clusterização** usando o algoritmo **K-Means**



# K-Means

É um método de **Clustering** que objetiva particionar **n** observações dentre **k** grupos onde cada observação pertence ao grupo mais próximo da média.

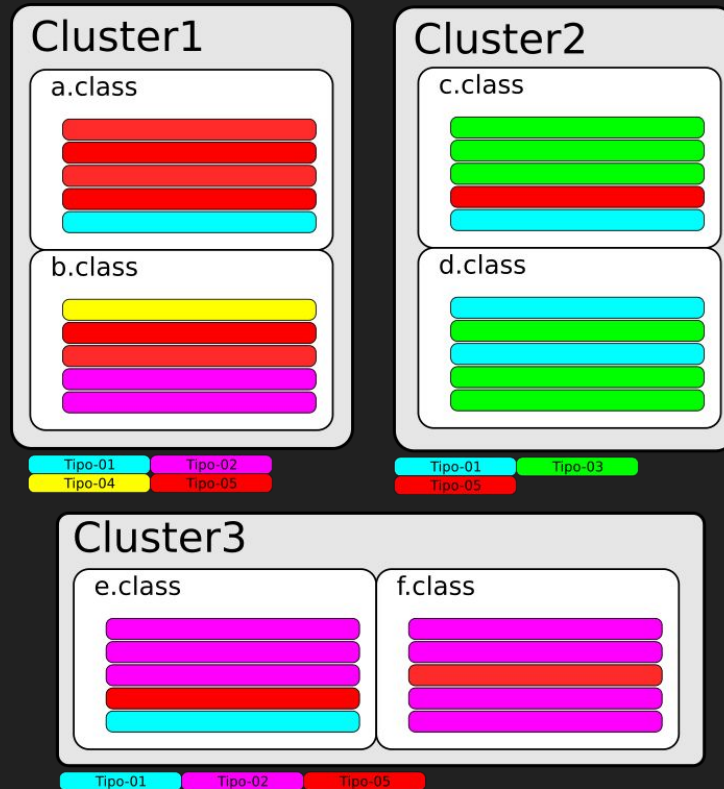
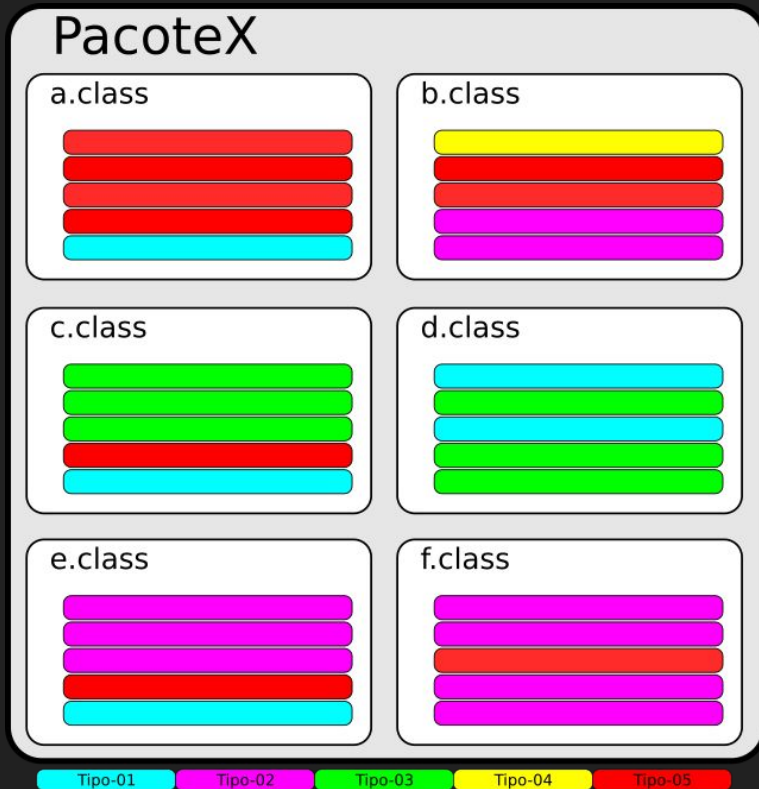
Isso resulta em uma divisão do espaço de dados em um Diagrama.



# Solução

- Eclipse AST
- Weka - Data Mining
- K-Means

# Solução - Após a Clusterização



# Vamos aos códigos



Obrigado

