Aula 2a Aprendizado de Máquina Conceitos fundamentais

Adaptado dos slides do Prof Samuel Botter

O que é Aprendizado de Máquina (Machine Learning)

- É uma subárea de Inteligência Artificial
- Envolve construir programas que aperfeiçoam-se automaticamente com sua experiência
- Existe na interseção da matemática e da estatística com a computação
- Enquanto a estatística concentra seu foco em descobrir algo interpretável a partir dos dados, o aprendizado de máquina tem por objetivo também transformar esses dados em uma ferramenta prática útil (um modelo).
- Estatística procura ensinar a humanos algo sobre os dados; Aprendizado de Máquina procura ensinar aos computadores algo sobre o mundo real.
- O aprendizado ocorre pela extração do máximo de informação possível dos dados, por meio de algoritmos que analisam a estrutura básica do conjunto de dados e buscam separar sinal (um padrão) de ruído (tudo o que não se encaixa no padrão).

Um pipeline genérico **Processamento** Exploração Iniciação do dataset Há relações relevantes entre as Ideia, Requisitos e Aquisição de Dados Pré-processamento e Seleção de variáveis? Características Por o modelo em produção e Calcular um índice de sucesso e Há muitos algoritmos e métodos expô-lo a dados reais interpretar os resultados disponíveis para escolher Desenvolvimento Avaliação Publicação do modelo Testing set Training set

Aprendizado de Máquina

Aprendizado supervisionado: treinar um modelo com dados de entrada e saída conhecidos a fim de pRedizer futuras saídas de dados novos

Aprendizado não-supervisionado: segmentar uma coleção de elementos com os mesmos atributos (agrupamento, clusterização)

Classificação

Regressão

Clusterização

SVM - Máquinas de Vetor de Suporte Regressão Linear

KNN - Vizinhos mais próximos

Métodos de montagem

Análise Discriminante Árvores de Decisão

Redes Neurais

Redes Neurais

K-means, k-medoids fuzzy C-means

Modelos ocultos de Markov

Mistura Gaussiana

Redes Neurais

Naïve BaSim

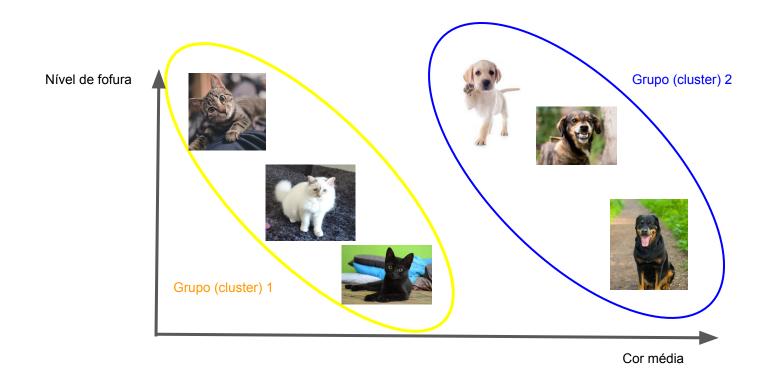
Tipos de algoritmos



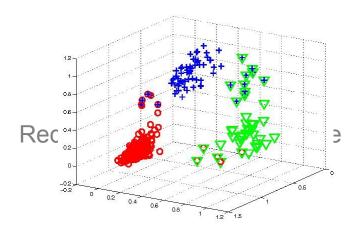




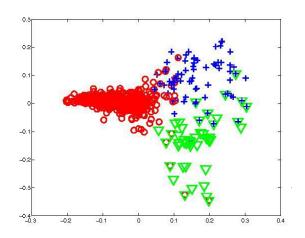
Aprendizado Não Supervisionado



Aprendizado Não Supervisionado

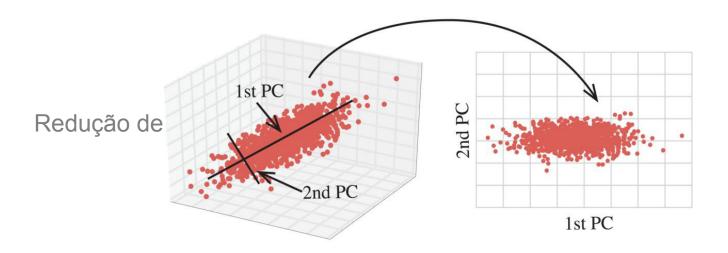


Espaço de características de maior dimensão



Espaço de características de menor dimensão

Aprendizado Não Supervisionado



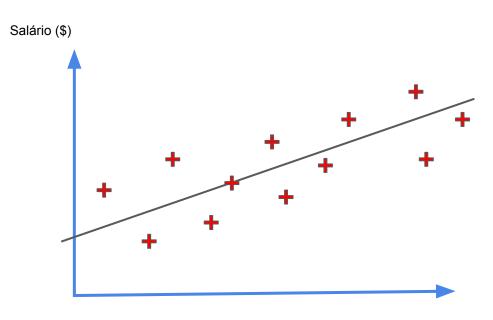
Espaço de características de maior dimensão

Espaço de características de menor dimensão

Aprendizado Supervisionado

Regressão Predição de variáveis numéricas

Tempo de experiência (meses)	Salário (R\$)
60	10.000,00
75	12.000,00
100	15.000,00



Experiência (t)