



Princípios do Machine Learning

- Objetivo: máquina olha os dados e aprender com eles.
- Exemplo: câmeras em aeroporto.

Figura 1 – Câmeras observando o espaço aéreo



Fonte: olaser/iStock.com.

Classificação

Figura 2 – Classificação do AM

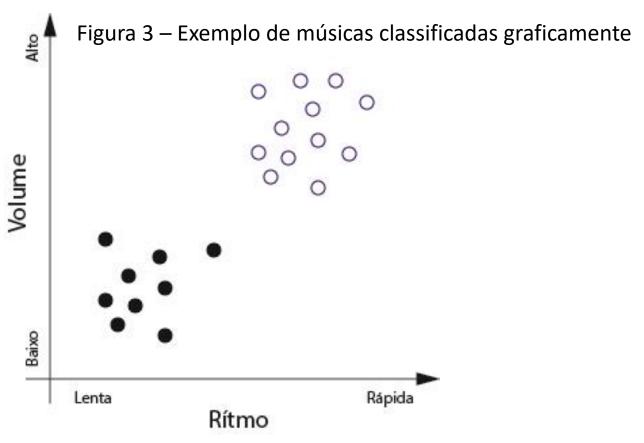


Fonte: elaborada pelo autor.

Existem vários outros tipos de classificações.

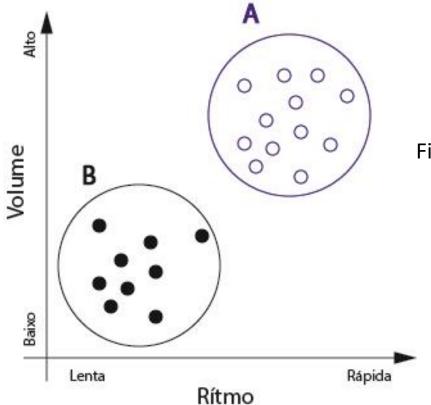
Aplicação de machine learning

Como a máquina pode tomar decisões?



Aplicação de machine learning

Como a máquina pode tomar decisões?

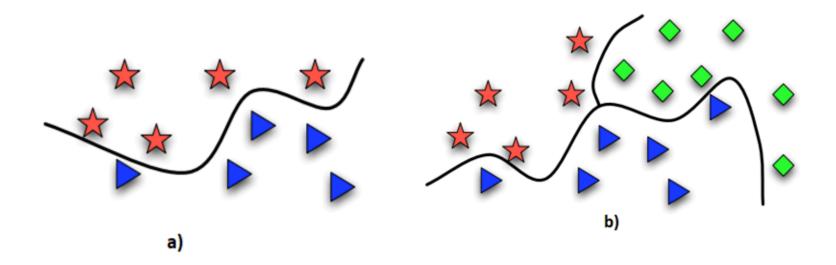


E um ponto fora da região?

Figura 4 – Exemplo de músicas classificadas graficamente (2)

Problemas clássicos do ML

Figura 5 – Classificações: a) binária e b) multiclasse

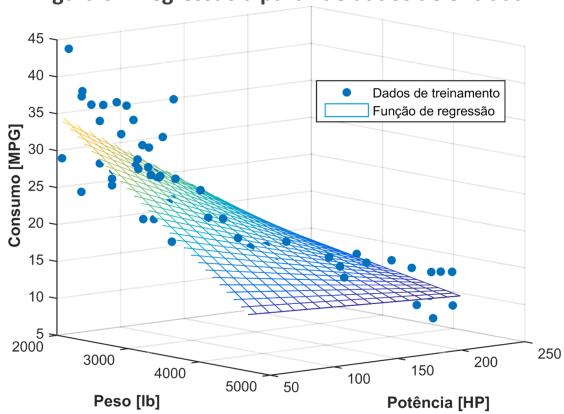


Fonte: Smola e Nishwanathan (2008, p. 10).

Problemas clássicos do ML

• Regressão.

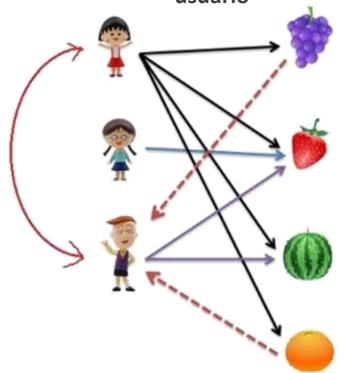
Figura 6 – Regressão a partir de dados de entrada



Problemas clássicos do ML

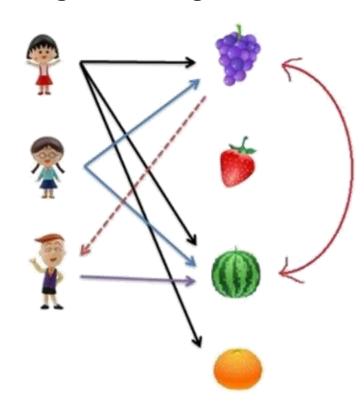
• Recomendação:

Figura 7 – Filtragem baseada em usuário



Fonte: Pinela (2017).

Figura 8 – Filtragem baseada em item



Fonte: Pinela (2017).



Classificação binária

- Não é tão simples quanto parece.
- \succ Estime $y \in \{\pm 1\}$ dado um padrão x contido em X.
- Complicações:
 - Aprendizado em lote: observa pares $(x_i, y_i) \in \{X, Y\}$, mas a estimação ocorre com X', que é desconhecido.
 - Estimativa com variáveis desconhecidas: ausência de informação sobre *X*.
 - ightharpoonup Correção de deslocamento por covariância: X e X' vem de fontes diferentes.



Complicações da classificação binária

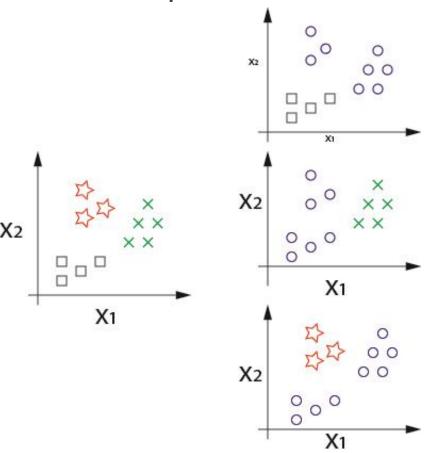
Aprendizado em lote.

Estimativa com variáveis desconhecidas.

Correção de deslocamento por covariância.

Classificação multiclasse One-vs-all

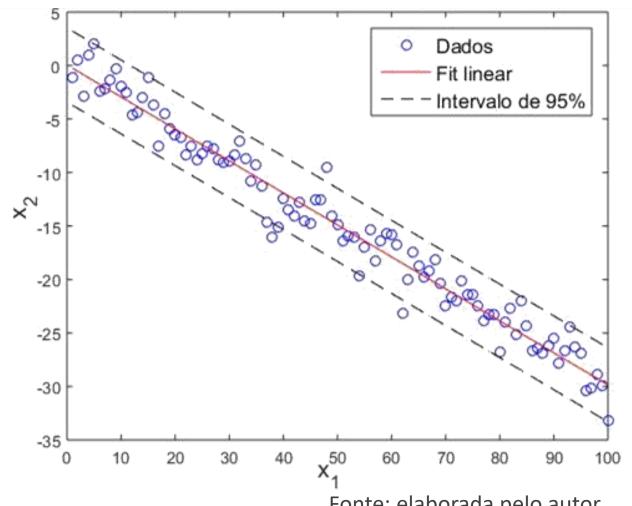
Figura 9 – Exemplo de classificação multiclasse do tipo *One versus all*



- É uma solução do aprendizado multiclasse.
- Objetivo: classificar a observação em alguma das classes existentes.
- Método: teste de todos os casos.

Regressão

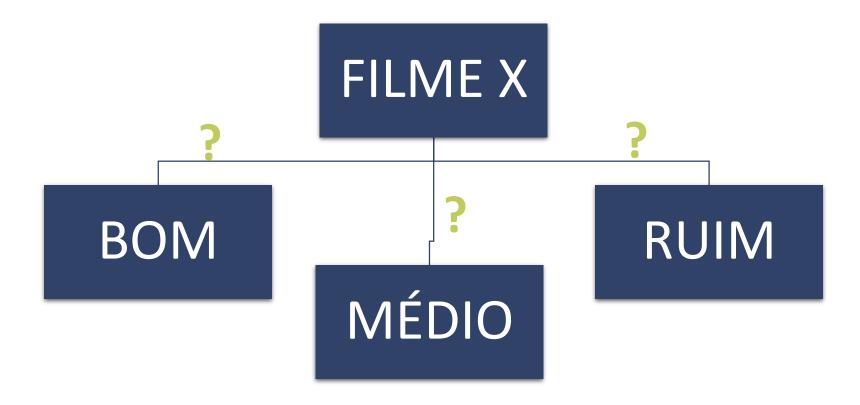
Figura 10 – Regressão linear com intervalo de predição de 95%





> Teoria em prática

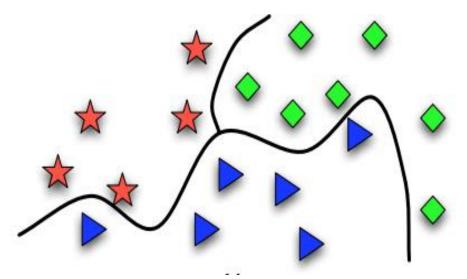
• Filmeflix quer separar os filmes de acordo com a notas?



Teoria em prática

- Problema de classificação multiclasse
- Variável estimada pode ter 3 classes: $y \in \{bom; médio; ruim\}$.
- O algoritmo deve conseguir separar os filmes nesses três grupos.

Figura 11 – Classificação multiclasse



Fonte: adaptada de Smola e Nishwanathan (2008, p. 10).



Como resolver?

Decision tree

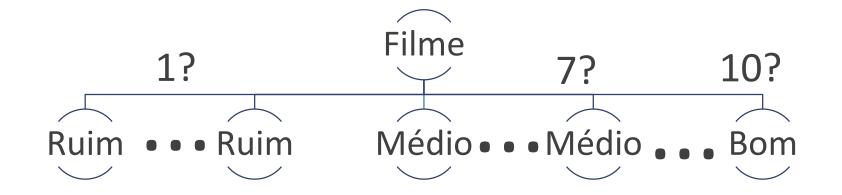
- Baseado em fluxos e testes.
- Nó representa um teste a ser feito.

Naive Bayes

- Algoritmo probabilístico.
- Baseado no teorema de Bayes.

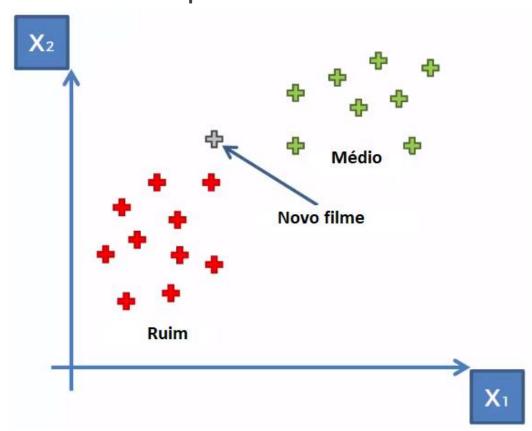
kNN

 K vizinhos próximos.



Como resolver?

Figura 12 – Problema de um novo dado inserido no conjunto para ser classificado





Dica do professor

• Familiarização com Python e R.





Fonte: adaptada de www.r-project.org. Acesso

em: 29 jul. 2019.

Fonte: adaptada de www.python.org. Acesso

em: 29 jul. 2019.

Referências

SMOLA, Alex; VISHWANATHAN, S. V. N. Introduction to Machine Learning. New York, NY: Cambridge University Press, 2008.

PINELA, Carlos. **Recommender Systems:** User-Based and Item-Based Collaborative Filtering. 2017. Disponível em:

https://medium.com/@cfpinela/recommender-systems-user-based-and-item-based-collaborative-filtering-5d5f375a127f>. Acesso em: 29 jul. 2019.

Referências

PYTHON CORE TEAM. **Python**: a dynamic, open source programming language. Python Software Foundation. Disponível em: https://www.python.org/. Acesso em: 29 jul. 2019.

R CORE TEAM. **R:** a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Disponível em http://www.R-project.org/. Acesso em: 29 jul. 2019.

