

# Cap. 01

## Aprendizagem Automática

IC 22-23

CPereira

### Machine Learning – O que é?

- Machine learning – Aprendizagem Automática
  - Extração de conhecimento a partir de dados;
  - Intersecção de diferentes áreas: Estatística, Inteligência Artificial, Ciências da computação
  - Ubíqua no dia a dia. Exemplos:
    - Recomendação de filmes ou séries a visualizar – Netflix
    - Quais os produtos a comprar – Amazon
    - Reconhecimento dos amigos em imagens – Facebook
    - Condução automática – Google
    - .....

# Machine Learning

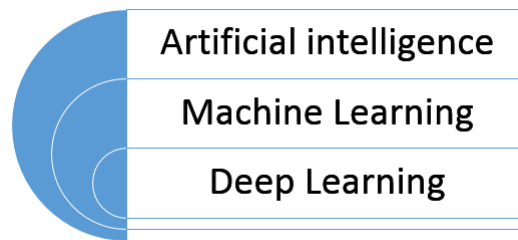
- Porquê aprender a partir de dados?
  - Tomar decisões com base em “handcoded rules” possui várias desvantagens:
    - Requer um conhecimento profundo do domínio do problema e a forma como o humano toma a decisão.
    - A lógica associada ao mecanismo de decisão é específica ao domínio do problema (filtro de spam). Como transferir conhecimento?
    - Como representar regras para o reconhecimento de pessoas em imagens?
      - A forma como uma imagem é representada em computador (conjunto de pixels) é muito diferente da forma como é percebida pelo humano!

# Machine Learning

- Que problemas pode resolver?
  - Todos os problemas que permitem automatizar o processo de decisão generalizando a partir de exemplos:
    - Reconhecimento de dígitos manuscritos (Mnist)
    - Classificação de imagens (médicas,...);
    - Detecção de fraudes em transações por cartão de crédito;
    - Etc...

# Machine Learning

- ...



# Machine Learning

- ...
  - Os exemplos anteriores assumem uma aprendizagem com base em exemplos classificados por humanos (aprendizagem supervisionada)
  - No modo não supervisionado, não existem “**labels**” associados aos **exemplos**:
    - Identificação de *tópicos* em “posts” de um blog;
    - **Segmentação** de clientes de acordo com as suas preferências (preferências similares devem pertencer ao mesmo grupo);
    - Detecção de **padrões** de acesso anormais a um “website”
    - ....

# Machine Learning

- Extracção de “Features”
  - Quais os dados a extrair?
  - Como construir uma boa representação dos dados?
  - Quais as características mais relevantes?

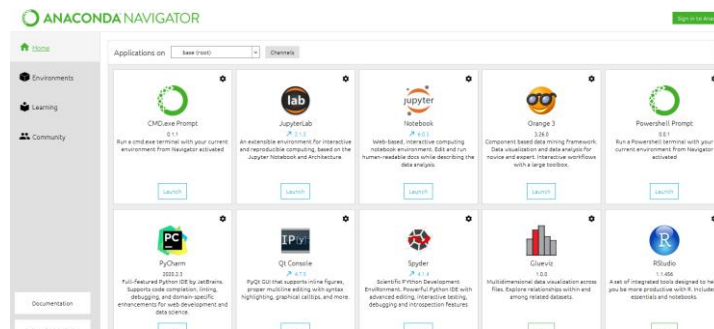


## Ferramentas computacionais

- Porquê Python?
  - Combina o poder computacional das linguagens de programação genéricas (C, C++,...) com a facilidade de utilização de linguagens script como Matlab ou R
  - Possui várias bibliotecas para leitura de dados, visualização, análise estatística, processamento de linguagem natural, imagem, texto, etc...
    - Instalar anaconda framework
      - Scikit-learn, Numpy, Scipy,...
      - Jupyter Notebook

# Ferramentas computacionais

- Anaconda Navigator



# Ferramentas computacionais

- Scikit-learn

## scikit-learn

Machine Learning in Python

Getting Started Release Highlights for 0.23 GitHub

- Simple and efficient tools for predictive data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

### Classification

Identifying which category an object belongs to.

**Applications:** Spam detection, image recognition, and more...

**Algorithms:** SVM, nearest neighbors, random forest, and more...

Examples

### Regression

Predicting a continuous-valued attribute associated with an object.

**Applications:** Drug response, Stock prices.

**Algorithms:** SVR, nearest neighbors, random forest, and more...

Examples

### Clustering

Automatic grouping of similar objects into sets.

**Applications:** Customer segmentation, Grouping experiment outcomes

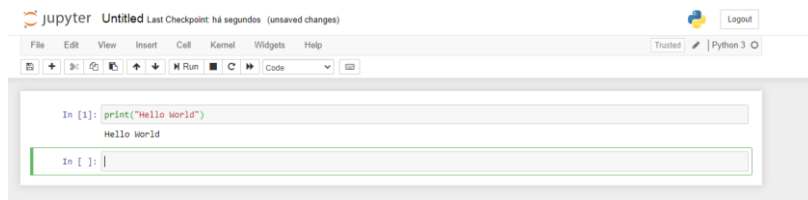
**Algorithms:** k-Means, spectral clustering, mean-shift, and more...

Examples

# Ferramentas computacionais

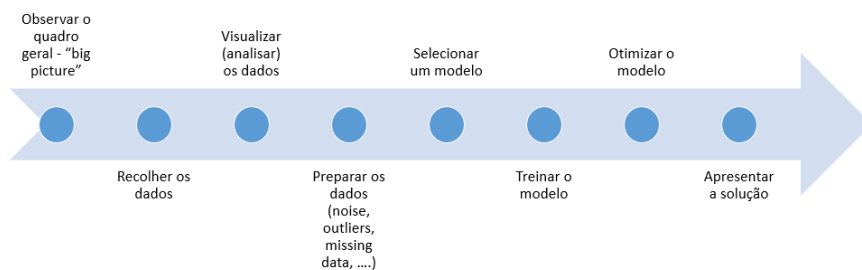
- Jupyter notebook

- Ambiente interativo para execução de código num browser
- Ideal para análise exploratória do problema
- Fácil a incorporação de código, texto e imagem...



## Primeiro exemplo

- Etapas de um processo de ML



## Primeiro Exemplo

- Iris dataset
  - O Problema
    - considerado o dataset "Hello world" da área de machine learning;
    - Distinguir três espécies de iris (setosa, versicolor, virginica)
    - Quais as características (features) dos dados?



## Primeiro Exemplo

- Conhecer o problema
  - Uma das etapas fundamentais do processo de ML é o conhecimento do problema e dados:
    - Que questão quer responder?
    - Tenho dados suficientes para representar o problema?
    - Qua características dos dados devo extrair?
    - Como vou avaliar o sucesso da minha solução?







## Primeiro exemplo

- Construir o modelo

```
In [6]: from sklearn import neighbors
classifier=neighbors.KNeighborsClassifier()
```



– <https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html>

## Primeiro exemplo

- Treinar o modelo

```
In [7]: classifier.fit(x_train,y_train)
```

- No scikit-learn, um estimador para classificação é um objeto Python que implementa os métodos `fit()` e `predict()`
  - <https://scikit-learn.org/stable/developers/develop.html>
  - <https://scikit-learn.org/stable/tutorial/basic/tutorial.html>

## Primeiro exemplo

- Avaliar o modelo

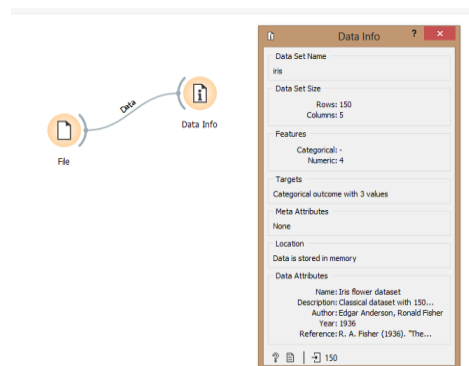
```
In [14]: predictions=classifier.predict(x_test)
         from sklearn.metrics import accuracy_score
         print(accuracy_score(y_test,predictions))

0.9866666666666667
```

- Se “normalize” == True, retorna a fração de amostras classificadas corretamente (float), caso contrário, retorna o número de amostras classificadas corretamente (int).
- [https://scikit-learn.org/stable/modules/model\\_evaluation.html](https://scikit-learn.org/stable/modules/model_evaluation.html)

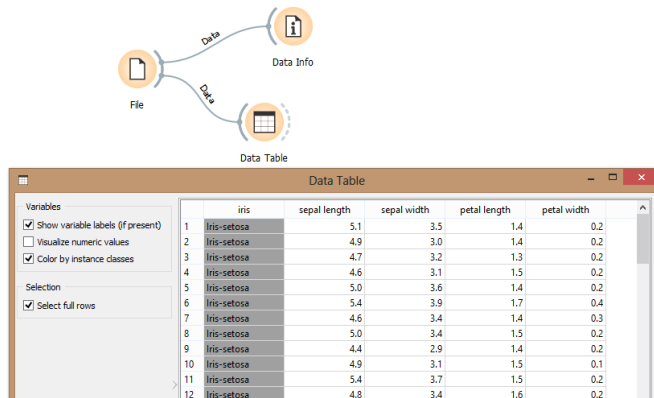
## Primeiro exemplo

- No Orange



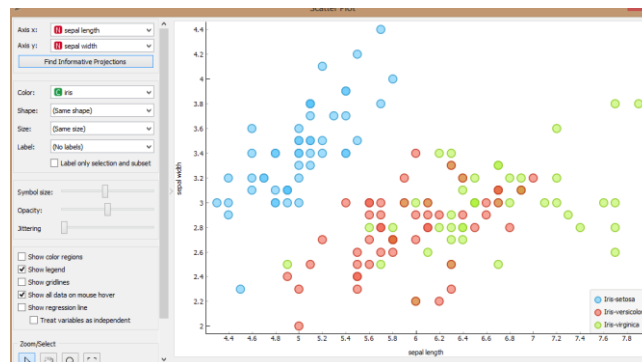
# Primeiro exemplo

• ...



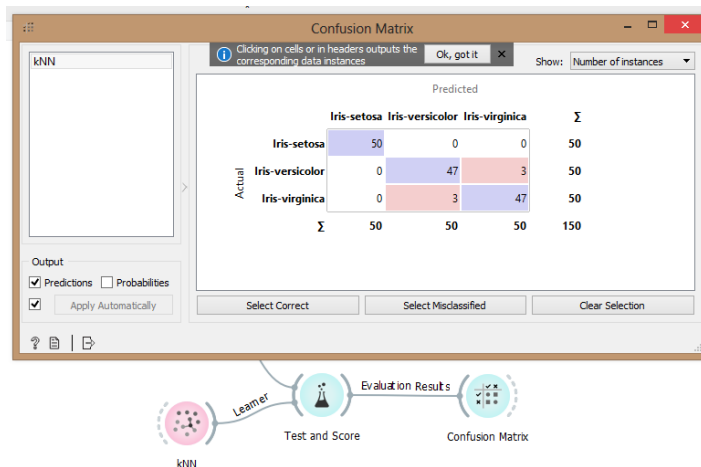
# Primeiro exemplo

• ...



## Primeiro exemplo

• ...



## Primeiro exemplo

- To do work...
  - Implementar o classificador para o problema “Wine Classification”
    - Matlab (live editor)
      - <https://www.mathworks.com/help/deeplearning/ug/wine-classification.html>
    - Quantas classes?
    - Quantas “features”?
    - Quantos exemplos?