

# Aula 1

---

quarta-feira, 8 de maio de 2024 14:54

1. Instalar o VirtualBox
2. Instalar o Ubuntu 22.04 ou 21.04 no VirtualBox
  - a. Se for visto que o Ubuntu está muito pesado, instalar o Xubuntu
3. Baixar o Jupyter Notebook (no Ubuntu)
4. Baixar as bibliotecas a serem utilizadas nas aulas
  - Numpy – biblioteca numérica
  - Pandas – biblioteca numpy
  - Seaborn - visualização de dados, contendo a matplotlib
  - Matplotlib - visualização de dados
  - Sci-py – fazer atividades e cálculos numéricos, parecido com numpy
  - Sklearn – biblioteca própria para ciência de dados, sci-kit learning
5. Importar as bibliotecas, sempre que criar um arquivo novo no Jupyter

Exemplo de como baixar as bibliotecas no Ubuntu

```
sudo aptitude install python3-numpy
```

```
sudo aptitude install python3-pandas
```

Exemplo de como importar as bibliotecas no arquivo no Jupyter

```
Import pandas as pd
```

## Aprendizado de Máquina

É um método de análise de dados que automatiza a construção de modelos analíticos. É um ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana.

Através de algoritmos para se criar um modelo (sistema), são inseridos dados para ir treinando esse modelo para chegar ao seu objetivo. Então cada vez mais são inserido dados e esse modelo é refinado e concertado em caso de erro.

Variável alvo: o que eu quero descobrir

### 1. Aprendizado Supervisionado

Por meio dos dados e do alvo (chamado de label em inglês) o modelo é treinado, a partir disso é possível usar esse modelo com outros tipos de dados.

Atributos/características (colunas)

- Variável dependente: alvo, depende da variável independente
- Variável independente

Modelos

- Classificação

Parâmetro de resultado: matriz de confusão

- Knn
- Arvore de decisão
- Svm (suporte vector machine)

- Regressão (adivinha o que vai dar na frente/futuro)

Parâmetro de resultado: desvio médio quadrático e desvio médio absoluto

- Polinomial
- linear

## 2. **Aprendizado Não Supervisionado**

Atributos/características (colunas)

- Variáveis independentes

Modelos

- Agrupamentos
- Pca (Análise do componente principal)
  - Reduz atributos(colunas), melhora o desempenho e diminui a complexidade

## 3. **Aprendizado por reforço**

Agente que tomam decisões, podendo ter reforço positivo ou negativo

# Python

## **Contas matemáticas**

É possível fazer contas matemáticas no prompt da maneira mais fácil, apenas inserindo a conta e executando

```
2+4
Ou também
print(2+4)
```

## **Criação de variáveis**

A variável é criada sem referenciar seu tipo, apenas seu nome e o seu valor. Não pode conter espaços no nome da variável e números no começo

```
valor = 2
operacao = (20+3)*1
```

## **Exibir no console/prompt**

Para exibir variáveis, resultados de operações, etc, pode ser escrito sem o print

```
valor          //2
2+4            //6
Ao invés de
printf(valor)
printf(2+4 )
```

## **Saber o tipo de uma variável**

```
print(type(valor))    //<class 'int'>
Ou
type(valor)           //int
```

### Alternando tipos de variáveis

- De float para int  
int(10.9)
- De int para float  
float(6)
- De float para round  
round(10.9)
- De string para int  
Int('38')

### Split de strings

```
elemento = "terra"
elemento[0:3]    //"ter"
elemento[0:1]    //"t"
```

### Somando e duplicando strings

```
elemento+elemento    //'terraterra'
Elemento*3            //'terraterterra'
```

### Inserindo dados com format

O método format() formata o valor especificado e os insere dentro do espaço reservado da string.

```
var = 'bebe'
'nome: {num}'.format(um=var)
Ou também
txt = "For only {price:.2f} dollars!"
txt.format(price = 49)
```

### Lista

As listas são arrays, podendo ser versáteis e conter int, string e float juntos

```
lista = [1, "dois", 3.2]
lista2 = [1, "dois", 3.2, [1, "dois", 3.2]]
lista[1]                                //"dois"
lista[0:2]                              //1,"dois"
```