

Aprendizado de Máquina e Deep Learning

Bibliotecas Python

Prof. Dr. Thiago Meirelles Ventura

Ambiente

- Linguagem de programação Python
- Uso do Google Colab
 - <https://research.google.com/colaboratory/>

Bibliotecas

- O Python possui diversas bibliotecas para auxiliar nos objetivos dos programas
 - Para ciência de dados não seria diferente
- Há três bibliotecas que podem ser muito úteis no carregamento, análise e processamento dos dados: Numpy, Matplotlib e Pandas
- Há bibliotecas para aplicar aprendizado de máquina

Numpy

- Pacote para realizar cálculos em Arrays Multidimensionais
 - Facilita o trabalho com listas e possui funções numéricas diversas

Numpy

script Numpy.ipynb

Matplotlib

- Facilita a criação de diferentes tipos de gráficos para auxiliar na visualização dos dados, que é importante para exploração, entendimento e demonstração de resultados

Matplotlib

script Matplotlib.ipynb

Pandas

- Pacote para facilitar a manipulação de dados, incluindo funções de manipulação de arquivos

Pandas

script Pandas.ipynb

Exercício 1

- Analise os dados do arquivo matogrosso.csv
 - Quais são as colunas?
 - Quais são os tipos de dados das colunas?
 - Quantos registros existem?
 - Há valores nulos? Quais?
 - Seria viável preencher os valores nulos? Se sim, como?

Exercício 2

- Visualize os dados do arquivo matogrosso.csv
 - Crie gráficos para apresentar as informações de
 - População estimada
 - Escolarização
 - Mortalidade infantil
 - PIB

Exercício 3

- Extraia informações do arquivo matogrosso.csv
 - Qual a média da área das cidades?
 - Qual a cidade com a menor população?
 - Qual a cidade com o maior IDHM?
 - Liste os prefeitos das cidades com PIB maior que 100.000
 - Crie uma nova coluna chamada SALDO que é o resultado de RECEITAS - DESPESAS