

Machine Learning com Python Básico

Prof. Marcelo Duduchi

1

Machine Learning com Python

EMENTA:

Aprendizagem, Agrupamento e Classificação, Python, Numpy, Pandas, Matplotlib, Scikit-Learn para Machine Learning, Agrupamentos e Classificação, Algoritmos K-means, Árvores de Decisão, KNN, SVM e Naive Bayes

OBJETIVO GERAL:

Desenvolver competências relacionadas à aplicação de algoritmos básicos de Machine Learning para agrupamento e classificação em Python.

OBJETIVOS ESPECÍFICOS:

- Desenvolver conceitos sobre aprendizagem supervisionada e não supervisionada
- Desenvolver competências no uso de Python básico para Machine Learning
- Desenvolver competências no uso de bibliotecas básicas para Machine Learning
- Desenvolver competências no uso de algoritmo para agrupamento
- Desenvolver competências no uso de algoritmos para classificação

2

Machine Learning com Python

CONTEÚDOS PROGRAMÁTICOS:

- Aprendizagem supervisionada e não supervisionada
- Agrupamento e classificação
- Python básico para Machine Learning
- Bibliotecas Numpy, Pandas, Matplotlib, Seaborn e Scikit-Learn
- Aplicação do algoritmo K-means
- Aplicação de algoritmos de Árvore de Decisão
- Aplicação do algoritmo KNN
- Aplicação do algoritmo SVM
- Aplicação do algoritmo Naive Bayes

3

Machine Learning com Python

CRONOGRAMA DE AULA:

- 1ª aula (4 horas): Conceitos básicos de IA e aprendizagem de máquina, uso do ambiente Colab e recursos de Python para Machine Learning.
- 2ª aula (4 horas): Bibliotecas de apoio a Machine Learning para computação científica e a recuperação, armazenamento, tratamento, armazenamento e visualização de dados.
- 3ª aula (4 horas): Agrupamento e classificação: uso dos algoritmos K-means e Árvore de Decisão.
- 4ª aula (4 horas): Algoritmos para classificação: uso dos algoritmos KNN, SVM e Naive Bayes.
- 5ª aula (4 horas): Algoritmos para classificação: uso prático dos algoritmos e comitês.

4

Bibliografia sobre IA

RUSSEL, S.; NORVIG, P. Inteligência Artificial. 3ed. Rio de Janeiro: Elsevier, 2013

<https://integrada.minhabiblioteca.com.br/#/books/9788521215479/cfi/4!/4/4@0.00:52.7>

FACIT, K.; LORENA, A. C.; GAMA, J.; ALMEIDA, T. A.; A. C. P. L. F., CARVALHO.

Inteligência Artificial: Uma abordagem de aprendizagem de máquina. 2ed. Rio de Janeiro: LTC, 2021

<https://integrada.minhabiblioteca.com.br/#/books/9788521637509/cfi/6/2!/4/2/2@0:66.9>

LUGER, George F. Inteligência artificial. 6ed. São Paulo: Pearson, 2013

<https://plataforma.bvirtual.com.br/Acervo/Publicacao/180430>

COPPIN, Ben. Inteligência artificial. Rio de Janeiro: LTC. 2013

<https://integrada.minhabiblioteca.com.br/#/books/978-85-216-2936-8>

SIMÕES, M. G.; SHAW, I. S. Controle e modelagem fuzzy. São Paulo: Blucher, 2007

<https://integrada.minhabiblioteca.com.br/#/books/9788521215479/cfi/4!/4/4@0.00:52.7>

5

Recursos

BANIN, S. L. Python 3: Conceitos e Aplicações. São Paulo: Érica, 2018

<https://integrada.minhabiblioteca.com.br/#/books/9788521215479/cfi/4!/4/4@0.00:52.7>

Google Colaboratory (ou apenas Colab) – Ambiente on-line de Notebooks Jupyter

Python para Ciência de Dados

<https://colab.research.google.com>

Biblioteca Pandas – Ferramenta para Ciência de Dados

<https://pandas.pydata.org/>

Biblioteca NumPy – Ferramenta de Computação Científica

<https://numpy.org/>

Biblioteca Matplotlib – Ferramenta para visualização de dados

<https://matplotlib.org/>

Biblioteca Seaborn – Ferramenta para estatística e visualização de dados

<https://seaborn.pydata.org/>

Biblioteca scikit-learn – Ferramenta para Análise de Dados Preditiva

<https://scikit-learn.org/>

6

Machine Learning com Python

Aula 02

7

Bibliotecas Numpy, Pandas e
Matplotlib

8

Bibliotecas Python para IA

- Estudaremos hoje algumas bibliotecas muito importantes para Inteligência Artificial. São elas:

- **Biblioteca NumPy** <https://numpy.org/>
- **Biblioteca Pandas** <https://pandas.org/>
- **Biblioteca Matplotlib** <https://matplotlib.org/>

9

Bibliotecas Python para IA

- Estudaremos hoje algumas bibliotecas muito importantes para Inteligência Artificial. São elas:

- **Biblioteca NumPy** <https://numpy.org/>
- **Biblioteca Pandas** <https://pandas.org/>
- **Biblioteca Matplotlib** <https://matplotlib.org/>

10

Numpy

- NumPy (Numerical Python) é uma biblioteca de código aberto criada em 2005 por Travis Oliphant para realizar operações em matrizes (arrays) multidimensionais.
- Implementa uma estrutura de dados denominada “ndarray” que permite a manipulação rápida e amigável com dados dispostos em matrizes, que não está disponível no Python.
- Usada por bibliotecas como: Pandas, Matplotlib, Scikit-Learn

11

Numpy

Recursos da biblioteca NumPy:

- Operações rápidas para tratamento e limpeza de dados
- Geração de subconjuntos e filtragens
- Estatística descritiva
- Manipulação de dados relacionais e dados em grupos
- funções matemáticas para operações rápidas com arrays, sem estruturas de repetição
- recursos de álgebra linear
- geração de números aleatórios

12

Numpy e Python

Diferenças entre “array” em Numpy e “list” em Python:

- Os arrays NumPy têm um tamanho fixo na criação, ao contrário das listas Python (podem crescer).
- Os elementos em uma matriz NumPy devem ser todos do mesmo tipo de dados e, portanto, terão o mesmo tamanho na memória.
- Os arrays NumPy facilitam operações matemáticas avançadas e outros tipos de operações em um grande número de dados.

13

Bibliotecas Pyton para IA

- Estudaremos hoje algumas bibliotecas muito importantes para Inteligência Artificial. São elas:
 - **Biblioteca NumPy** <https://numpy.org/>
 - **Biblioteca Pandas** <https://pandas.org/>
 - **Biblioteca Matplotlib** <https://matplotlib.org/>

14

Pandas

- Pandas é uma biblioteca de código aberto criada em 2008 por Wes McKinney para manipulação e análise de dados.
- É uma biblioteca Python que fornece uma série de recursos de leitura e manipulação de grandes bases de dados com uma grande variedade de comandos
- Permite abrir e manipular bases de dados (datasets) em Excel (.xlsx) e CSV (.csv) criando os objetos chamados “DataFrames” que podem ser manipulados.

15

Pandas

Recursos da biblioteca Pandas:

- Ferramentas para ler e escrever dados entre diferentes estruturas de dados e formatos de arquivo
- Alinhamento de dados e manipulação de dados ausentes
- Reformatação de matrizes (dados)
- Inclusão e deleção de colunas em conjuntos de dados
- Ferramentas para fundir (merging) ou juntar (join) conjuntos de dados
- Filtro e limpeza de dados

16

Bibliotecas Python para IA

- Estudaremos hoje algumas bibliotecas muito importantes para Inteligência Artificial. São elas:

- **Biblioteca NumPy** <https://numpy.org/>
- **Biblioteca Pandas** <https://pandas.org/>
- **Biblioteca Matplotlib** <https://matplotlib.org/>

17

Matplotlib

- A Matplotlib é uma biblioteca abrangente para criar visualizações estáticas, animadas e interativas em Python
- Foi criada originalmente por John D. Hunt (biólogo) em 2003 que tem desenvolvimento comunitário liderado por Michael Droettboom.
- Segundo o próprio site a Matplotlib torna as coisas fáceis, mais fáceis ainda e as difíceis possíveis.

18

Matplotlib

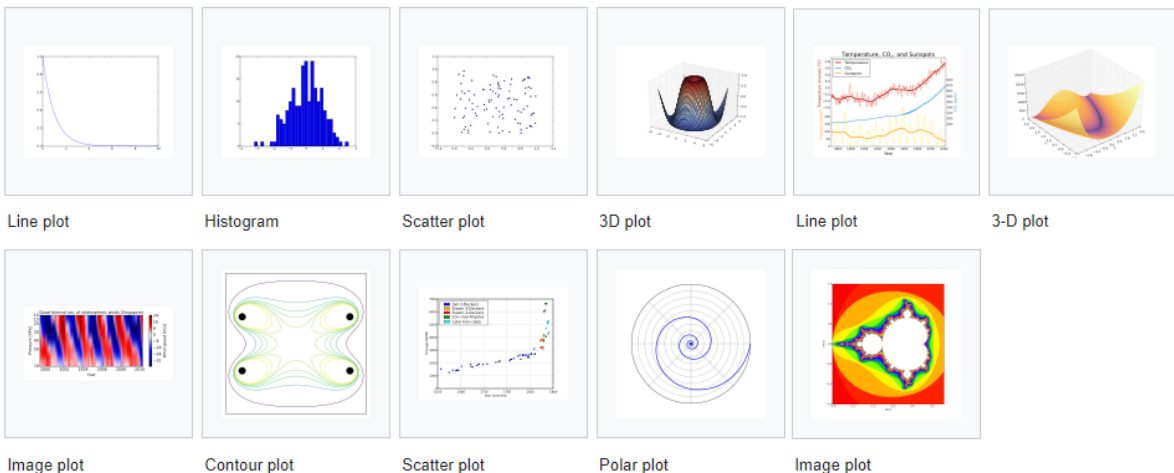
Recursos da biblioteca Matplotlib:

- Criação de gráficos em qualidade de publicação
- Criação de figuras interativas que podem ser ampliadas, deslocadas e atualizadas
- Personalização o estilo visual e o layout
- Exportação para vários formatos de arquivo

19

Matplotlib

Recursos da biblioteca Matplotlib:



20

NumPy, Pandas e Matplotlib

- Conheceremos as bibliotecas NumPy, Pandas e Matplotlib por meio de três roteiros que depois você deve colocar em notebooks para experimentar seus recursos.
- Em alguns casos usaremos bases de dados (datasets) para ilustrar manipulações e apresentações de dados possíveis.

21

Carregando bibliotecas em Python

Para usar essas bibliotecas no Colab basta fazer o “import” porque elas já estão instaladas:

```
import <nome> as <apelido>
```

Note que ao importá-las podemos dar apelidos para elas facilitando o uso. Este recurso é muito utilizado porque diminui muito a carga de digitação.

22

Carregando bibliotecas em Python

Exemplos:

```
import numpy as np
```

```
import pandas as pd
```

```
import matplotlib as mpl
```

```
import matplotlib.pyplot as plt
```

23

Conhecendo as bibliotecas

Executaremos agora nossos notebooks...

24