
ESPECIALIZAÇÃO EM CIÊNCIA DE DADOS – PUC-RIO

MACHINE LEARNING

AULA 0: APRESENTAÇÃO DA DISCIPLINA

Tatiana Escovedo, PhD.
tatiana@inf.puc-rio.br

EMENTA RESUMIDA

- Algoritmos de Machine Learning;
- Aprendizado supervisionado e não-supervisionado;
- Problemas de Classificação;
- Problemas de Regressão;
- Problemas de Associação;
- Problemas de Agrupamento;
- Modelagem;
- Métricas de Avaliação;
- Comparação de Modelos.
- **NÃO É UM CURSO DE PYTHON!**



SOFTWARES PARA OS LABORATÓRIOS

Linguagem

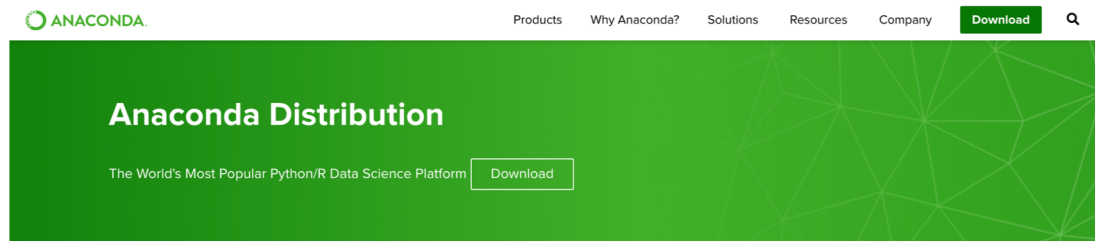
- Python
 - 😊 É fácil de aprender
 - 😊 É muito popular
 - 😊 Linguagem de uso geral
 - 😊 Linguagem interpretada (maior facilidade)
 - 😊 Grande comunidade
 - 😞 Mais lenta do que linguagens não-interpretadas, como Java.
 - 😞 Não é ideal para aplicações com concorrência.

Ferramentas

- Anaconda Python OU
- Jupyter Notebook OU
- Jupyter Lab OU
- Google Colab: <https://colab.research.google.com>

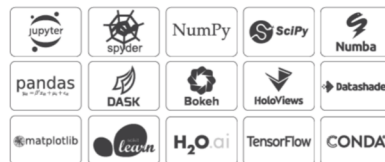
PREPARAÇÃO DO AMBIENTE DE DESENVOLVIMENTO

- Instalar o Anaconda: www.anaconda.org



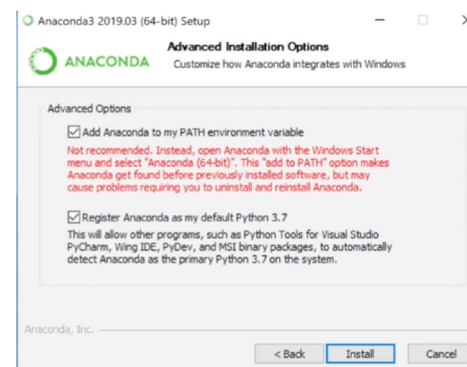
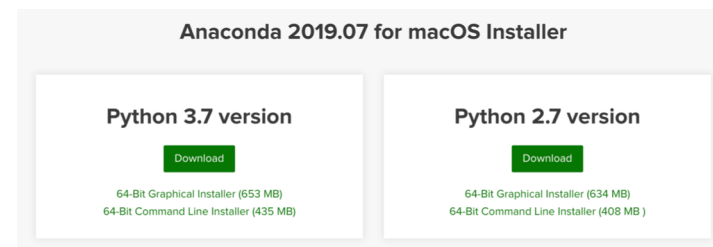
The open-source Anaconda Distribution is the easiest way to perform Python/R data science and machine learning on Linux, Windows, and Mac OS X. With over 15 million users worldwide, it is the industry standard for developing, testing, and training on a single machine, enabling *individual data scientists* to:

- Quickly download 1,500+ Python/R data science packages
- Manage libraries, dependencies, and environments with Conda
- Develop and train machine learning and deep learning models with scikit-learn, TensorFlow, and Theano
- Analyze data with scalability and performance with Dask, NumPy, pandas, and Numba
- Visualize results with Matplotlib, Bokeh, Datashader, and Holoviews



Windows | macOS | Linux

- Instalar o Python 3+



AVALIAÇÃO

- I prova (NP) e I trabalho em grupo de 3 a 5 alunos (NT)
 - Se $NP < 5 \rightarrow$ Prova Final (substitui a NP)
- Média Final (MF): $(NP + NT)/2$
 - Se $MF \geq 6 \rightarrow$ Aprovado

BIBLIOGRAFIA RECOMENDADA

Livros

- **Data Science do Zero. Primeiras Regras com o Python** - Joel Grus
- **Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications** – Laura Igual and Santi Seguí

Cursos

- Python Fundamentos para Análise de Dados - <https://www.datascienceacademy.com.br/pages/curso-python-fundamentos-para-analise-de-dados> (gratuito)
- Curso rápido de Python do zero: <https://medium.com/trainingcenter/aprendendo-python-do-zero-b9a13d8646b3>
- Machine Learning Crash Course — <https://developers.google.com/machine-learning/crash-course/> (gratuito)

- Learn Python (Code Academy) — <https://external.codecademy.com/learn/learn-python>
- Python For Beginners — <https://external.codecademy.com/learn/learn-python-3>

Blogs

- <https://machinelearningmastery.com/>
- <https://www.kdnuggets.com>
- <http://datascienceacademy.com.br/blog>
- Checklist de projetos para Machine Learning: <https://medium.com/databootcamp/meu-checklist-de-projetos-de-aprendizado-de-m%C3%A1quina-34328850d7ab>
- Cheatsheets for Data Science: <https://startupsventurecapital.com/essential-cheat-sheets-for-machine-learning-and-deep-learning-researchers-efb6a8ebd2e5>
- Exemplos em Python: <https://github.com/donnemartin/data-science-python-notebooks>