



EEL891 – Aprendizado de Máquina



Heraldo Luís Silveira de Almeida, D.Sc.

Professor Associado

Depto. Engenharia Eletrônica e de Computação
Universidade Federal do Rio de Janeiro (UFRJ)

Sala H219 Gabinete 18

Tel. 55 21 3938-8192

www.del.ufrj.br/~heraldo

heraldo@ufrj.br



0

Informações sobre a Disciplina



0. Informações sobre a Disciplina

0.1. Objetivo

0.2. Metodologia

0.3. Ementa

0.4. Bibliografia

0.5. Avaliação

0.6. Links Úteis

0.7. Uma Enquete Rápida



0.1. Objetivo

Proporcionar ao aluno conhecimentos na área de **Machine Learning**, capacitando-o a construir e validar modelos para tarefas tais como:

- ☐ classificação
- ☐ regressão
- ☐ agrupamento (*clustering*)
- ☐ detecção de anomalias
- ☐ separação de sinais
- ☐ otimização de processos decisórios

aplicando os mais modernos e eficazes algoritmos de aprendizado automático com base em dados.



0.2. Metodologia

- ☐ Exposição teórica dos conceitos matemáticos, técnicas e algoritmos de aprendizado de máquina.
- ☐ Experimentos computacionais apresentados pelo professor em sala de aula.
- ☐ Uso de linguagem de programação e ferramentas de software gratuitas amplamente utilizadas no mercado.
- ☐ Serão utilizadas neste semestre a linguagem **Python** e as ferramentas **scikit-learn** e **TensorFlow**.



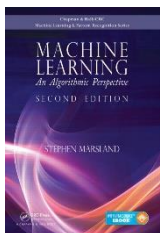


0.3. Ementa

- ❑ Aprendizado supervisionado: algoritmos para classificação e regressão (perceptron, modelos *bayesianos*, redes neurais, SVM, k-NN, árvores/florestas de decisão, etc.).
- ❑ Generalização, medidas de erro, treinamento e teste, viés e variância, *overfitting*, técnicas de regularização e algoritmos de validação.
- ❑ Aprendizado não-supervisionado: algoritmos para agrupamento, detecção de anomalia, separação de sinais e estimação de densidade.
- ❑ Aprendizado por reforço: modelagem como processo de decisão de Markov e algoritmos de otimização de estratégia de decisão.
- ❑ Redes neurais profundas e *deep learning*.

0.4. Bibliografia

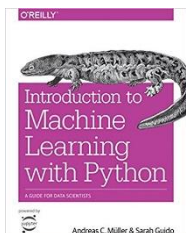
Livros nos quais o curso se baseia:



MARSLAND, S.

Machine Learning – An Algorithmic Perspective.

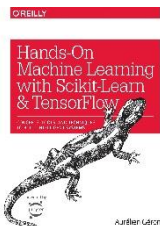
CRC Press, 2015.



MÜLLER, A. C. & GUIDO, S.

Introduction to Machine Learning with Python.

O'Reilly Media, 2017.



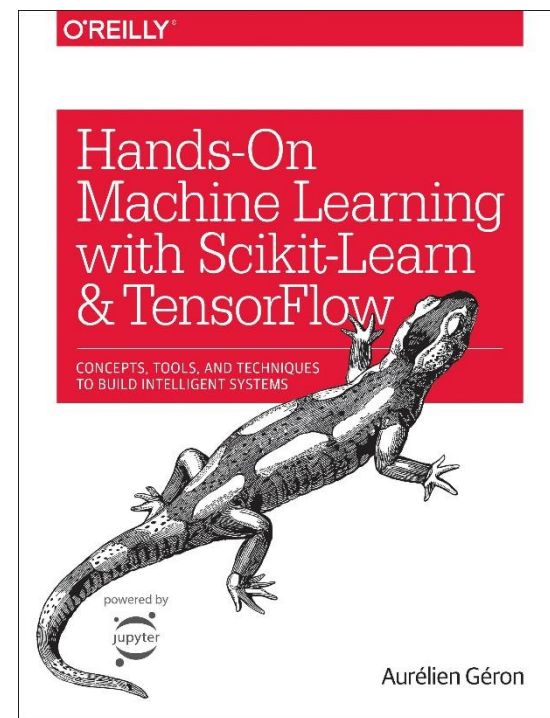
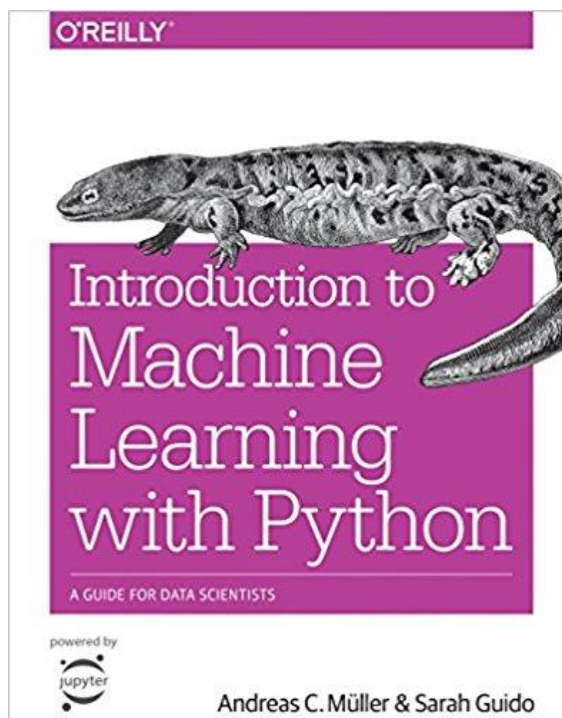
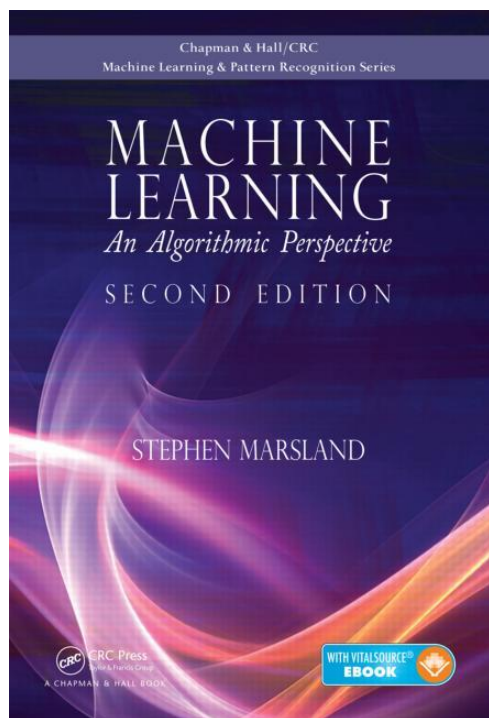
GÉRON, A.

Hands-On Machine Learning with Scikit-Learn and TensorFlow.

O'Reilly Media, 2017.

0.4. Bibliografia

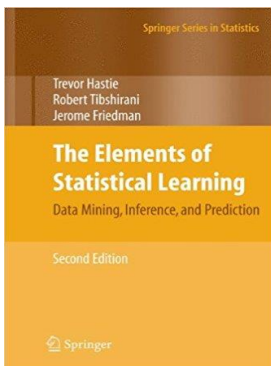
Livros nos quais o curso se baseia:



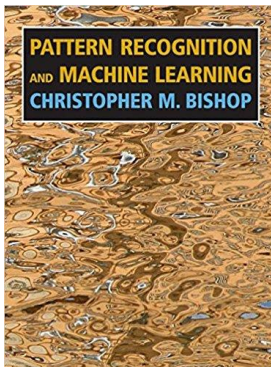
0.4. Bibliografia

Outros livros:

Clássicos sobre o assunto:

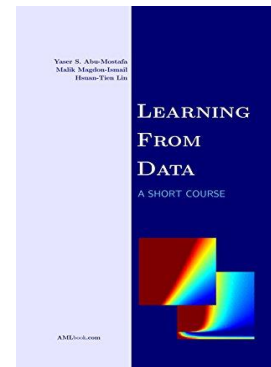


BISHOP, C. M.
*Pattern Recognition
and Machine Learning.*
Springer, 2013.

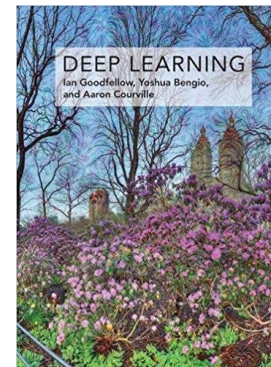


HASTIE, T.,
TIBSHIRANI, R.
& FRIEDMAN J.
*The Elements of
Statistical Learning.*
Springer, 2016.

Outros recomendados:



ABU-MOSTAFA, Y. S.,
MAGDON-ISMAIL, M.
& LIN, H.
Learning from Data.
AMLbook, 2012.



GOODFELLOW, I.,
YOSHUA, B.
& COURVILLE, A.
Deep Learning
The MIT Press, 2016.



0.5. Avaliação

- ☐ Trabalhos práticos individuais envolvendo linguagem Python e bibliotecas Scikit-Learn e/ou TensorFlow.
- ☐ Quantidade de trabalhos ainda a ser definida (provavelmente entre 2 e 4 trabalhos).
- ☐ Pontuação dos trabalhos na escala de 0 a 10:
 - 5 pontos pela entrega do trabalho (código-fonte + dados + relatório em formato PDF) conforme especificado;
 - 2 pontos pela qualidade do relatório em formato PDF;
 - 3 pontos pelo desempenho da máquina desenvolvida;
 - “multas” por atraso na entrega a serem definidas.
- ☐ Média Final = média aritmética das notas dos trabalhos



0.6. Links Úteis

Pacotes de Software Utilizados no Curso:

- ☐ Python (<https://www.python.org>)
- ☐ pandas (<https://pandas.pydata.org>)
- ☐ NumPy (www.numpy.org)
- ☐ Matplotlib (<https://matplotlib.org>)
- ☐ scikit-learn (<http://scikit-learn.org>)
- ☐ TensorFlow (<https://www.tensorflow.org>)

Roteiro de Instalação de Software para o Curso:

- ☐ Linux Ubuntu (www.del.ufrj.br/~heraldo/eel891/RoteiroUbuntu.pdf)



0.6. Links Úteis

Bons repositórios de dados para usar nos experimentos:

- ☐ UCI Machine Learning Repository
<http://archive.ics.uci.edu/ml/index.php>
- ☐ Kaggle Datasets
<https://www.kaggle.com/datasets>
- ☐ OpenML
<https://www.openml.org/>

Sites que listam diversos repositórios de dados para ML:

- ☐ KD Nuggets
<https://www.kdnuggets.com/datasets/index.html>
- ☐ Wikipedia
https://en.wikipedia.org/wiki/List_of_datasets_for_machine_learning_research



0.6. Links Úteis

**Cursos Online Gratuitos de ML de Universidades dos EUA
(provavelmente melhores que o meu!) :**

- ☐ Prof. Andrew Ng (Stanford University)
<https://www.coursera.org/learn/machine-learning>
- ☐ Prof. Yaser Abu-Mostafa (CalTech University)
<https://www.coursera.org/learn/machine-learning>
- ☐ Prof. John Paisley (Columbia University)
<https://www.edx.org/course/machine-learning-columbiacx-csmm-102x-3>



0.6. Links Úteis

Tarefas que todos devem fazer:

1. Cadastrar-se no Kaggle (www.kaggle.com)
2. Instalar o Anaconda (www.anaconda.org)



0.7. Uma Enquete Rápida

- ☐ Quantos de vocês já **ouviram falar** de Machine Learning ?
- ☐ Quantos de vocês **estudaram algo** sobre Machine Learning?
- ☐ Quantos de vocês já **implementaram** Machine Learning ?
- ☐ Quantos de vocês já usaram **scikit-learn** ?
- ☐ Quantos de vocês já usaram **TensorFlow** ?