DataRoot | Data Science Basics | Part 2

(Machine Learning)

**Necessary:**

**Main algorithms/methods:**

**Gradient Descent**

**Linear regression**

**Logistic regression**

**Maximum likelihood estimation**

**Generalized linear models**

**Naive Bayes Classifier**

**Gaussian Discriminant Analysis**

**Regularization**

**SVM. Kernels**

**K-means**

**PCA**

**Random Forest\*\*\***

**Neural Networks (**[**brief overview**](https://www.youtube.com/playlist?list=PLiaHhY2iBX9hdHaRr6b7XevZtgZRa1PoU)**)**

**Learn, implement, test, compare.**

**Choose any dataset here:** [**http://www.datapure.co/open-data-sets**](http://www.datapure.co/open-data-sets)

**So you need to:**

1. **send implemented algorithms/methods (simple and working ones)**
2. **try it on any dataset you like and compare with sklearn realization (for models)**
3. **talk about them on the next meeting**

[**http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/**](http://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/) **- systematization**

[**http://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html**](http://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html) **- must know**

**Data Mining Video Lectures: (quick overview)**

[**https://www.youtube.com/playlist?list=PLLssT5z\_DsK9JDLcT8T62VtzwyW9LNepV**](https://www.youtube.com/playlist?list=PLLssT5z_DsK9JDLcT8T62VtzwyW9LNepV)

**Machine Learning Coursera Course:**

[**https://www.coursera.org/learn/machine-learning**](https://www.coursera.org/learn/machine-learning)

**Additional books:**

1. Hastie, Tibshirani, and Friedman's [The Elements of Statistical Learning](http://statweb.stanford.edu/~tibs/ElemStatLearn/)

2. Bishop's [Pattern Recognition and Machine Learning](http://research.microsoft.com/en-us/um/people/cmbishop/prml/)

3. David Barber's [Bayesian Reasoning and Machine Learning](http://web4.cs.ucl.ac.uk/staff/D.Barber/pmwiki/pmwiki.php?n=Brml.HomePage)

4. (Great Book)Kevin Murphy's [Machine learning: a Probabilistic Perspective](http://www.cs.ubc.ca/~murphyk/MLbook/)

5. Foundations of Machine Learning, [Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar](http://www.amazon.com/dp/026201825X)

6. Learning From Data,[Yaser S. Abu-Mostafa, Malik Magdon-Ismail, Hsuan-Tien Lin](http://www.amazon.com/gp/product/1600490069)

7. Information Theory, Inference, and Learning Algorithms,[David J. C. MacKay](http://www.amazon.com/Information-Theory-Inference-Learning-Algorithms/dp/0521642981)[free pdf]

8. All of Statistics,[Larry Wasserman](http://www.amazon.com/All-Statistics-Statistical-Inference-Springer/dp/0387402721)

9. Probabilistic Graphical Models: Principles and Techniques, [Daphne Koller, Nir Friedman](http://www.amazon.com/Probabilistic-Graphical-Models-Principles-Computation/dp/0262013193)

10. Gaussian Processes For Machine Learning,[Carl Edward Rasmussen, Christopher K. I. Williams](http://www.amazon.com/Gaussian-Processes-Learning-Adaptive-Computation/dp/026218253X) [free pdf]

11. [Machine Learning with R](http://bit.ly/16bxAum)

12. [Building Machine Learning Systems with Python](http://bit.ly/1vHzOWS)

13. [Machine Learning with Spark](http://bit.ly/1GnnQdE)

14. [Matrix Computations (Johns Hopkins Studies in the Mathematical Sciences): Gene H. Golub, Charles F. Van Loan: 9781421407944: Amazon.com: Books](http://www.amazon.com/Computations-Hopkins-Studies-Mathematical-Sciences/dp/1421407949/ref=sr_1_1?ie=UTF8&qid=1456795538&sr=8-1&keywords=matrix+computations)

15. [Amazon.com: Convex Optimization (9780521833783): Stephen Boyd, Lieven Vandenberghe: Books](http://www.amazon.com/Convex-Optimization-Stephen-Boyd/dp/0521833787/ref=sr_1_1?ie=UTF8&qid=1456795591&sr=8-1&keywords=convex+optimization)