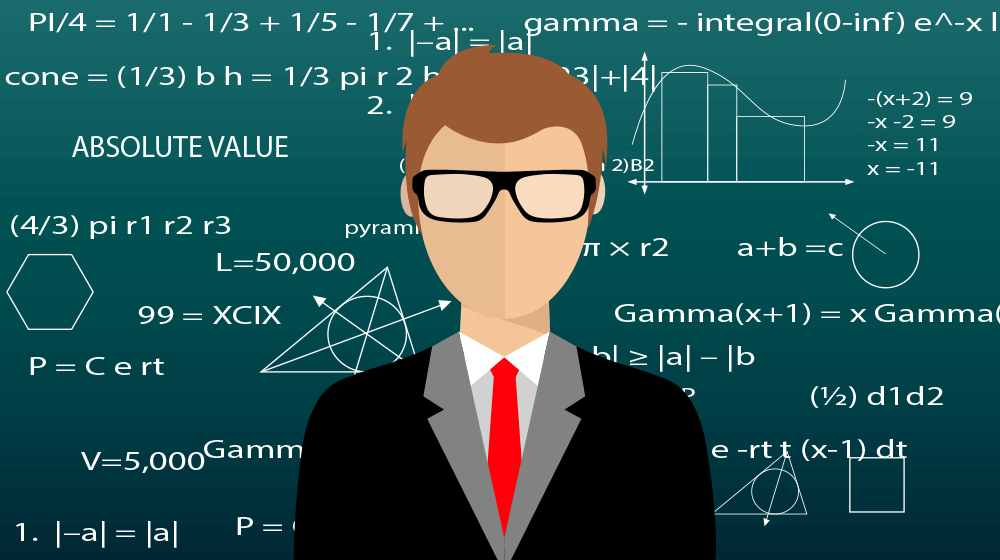
Data Scientist Learning Path 2019

[](https://medium.com/@ashishpatel.ce.2011?source=post_header_lockup)

[Ashish Patel](https://medium.com/@ashishpatel.ce.2011)Follow

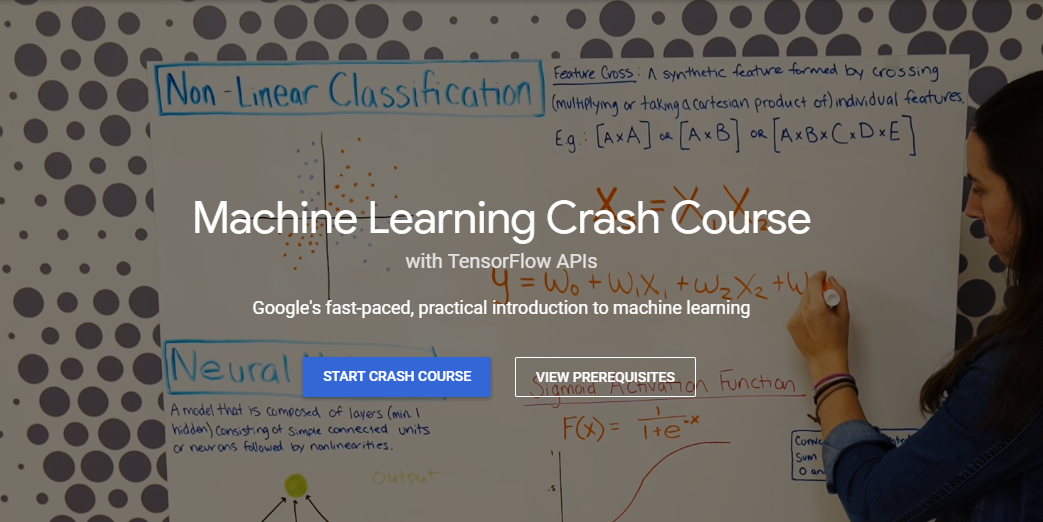
Jul 17, 2018

Machine Learning Series!!!



**Hi**Folks, So many have asked me basic question always, **“Can you please suggest me the best path for become data scientist ? ”.** When I was beginner so many website help me to learn the data science.Special thanks to [analyticsvidhya.com](http://analyticsvidhya.com/) , [AndrewNG Course](https://www.coursera.org/learn/machine-learning" \t "_blank), and [Jason Brownlee](https://machinelearningmastery.com/)which made me give good insight of data science.So, Today I am sharing here to learning path for become to better data scientist.

**M**ust Suggested Special Course of Machine Learning by Google [Click Here](https://developers.google.com/machine-learning/crash-course/)



**Step 1 : Basic Python Learning**

Choose one language for machine learning is must necessary and I suggest python because it is most popular language in Data scientist.you can also go with R.But I suggest python to learn don’t worry.I am sharing here both language learning resources.

**Python Resources to Learn**

1. [Learn Python for Data Science — Online Course | DataCamp](https://www.datacamp.com/courses/intro-to-python-for-data-science)
2. [Basic Python by Corey Schaftr](https://www.youtube.com/playlist?list=PL-osiE80TeTskrapNbzXhwoFUiLCjGgY7)
3. **Books (mandatory)** — [Python for Data Analysis](http://www3.canisius.edu/~yany/python/Python4DataAnalysis.pdf) — This book covers various aspects of Data Science including loading data to manipulating, processing, cleaning and visualizing data. Must keep reference guide for Pandas users.

**R Resources to Learn**

1. [R Programming by Newboston](https://www.youtube.com/watch?v=X67No4239Ys&list=PL6gx4Cwl9DGCzVMGCPi1kwvABu7eWv08P)
2. [R by Tutorialspoint.com](https://www.tutorialspoint.com/r/index.htm)
3. [R By Datacamp](https://www.datacamp.com/courses/free-introduction-to-r)
4. **Books** — [R for Data Science](http://r4ds.had.co.nz/) — This is your one stop solution for referencing basic materials on R.

**Step 2 : Develop Skills in Algebra, Statistics, and ML**

Data scientists are better at statistics than any software engineer and are better at software engineering than any statistician. Our idea is to maintain the right balance and avoid too much or not enough to emphasize either.

**Descriptive Statistics — 1 week**

* **Course (mandatory) —**[Descriptive Statistics from Udacity](https://www.udacity.com/course/intro-to-descriptive-statistics--ud827) is a basic and must do course to get started.
* **Books (optional)**— Supplement your online course with [online stats book](http://onlinestatbook.com/2/index.html). A good book for any one looking for learning basic statistics.

**Probability — 2 weeks**

* **Course (mandatory)** — [Introduction to probability — The science of uncertainty](https://www.edx.org/course/introduction-probability-science-mitx-6-041x-2) is an excellent course on edX to learn concepts of probability like conditional probability and probability distributions.
* **Books (optional)** — The textbook [Introduction to probability — Berkley’s stats 134 standard textbook](http://www.stat.berkeley.edu/~aldous/134/grinstead.pdf) will supplement the course above and can be used as a good reference material.

**Inferential Statistics — 2 weeks**

* **Course (mandatory)**— [Intro to Inferential Statistics from Udacity](https://www.udacity.com/course/intro-to-inferential-statistics--ud201) — Once you have gone through the descriptive statistics course, this course will take you through statistical modeling techniques and advanced statistics.
* **Books (optional)** — [Online Stats Book](http://onlinestatbook.com/2/index.html) — This online book can be used for a quick reference for inference tasks.

**Linear Algebra — 1 week**

* **Course (mandatory)**
* [Linear Algebra — Khan Academy](https://www.khanacademy.org/math/linear-algebra) : This concise and an excellent course on Khan Academy will equip you with the skills necessary for Data Science and Machine Learning.

**Books (optional)**

* [Linear Algebra/ Levandosky](https://www.amazon.com/gp/product/0536667470/ref=as_li_qf_sp_asin_il_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=0536667470&linkCode=as2&tag=theopesoudats-20&linkId=YLLIM4C4LJSOZIQR) — This is an often cited book to Stanford graduates for Linear Algebra.
* [The Manga guide to Linear Algebra](https://www.amazon.com/Manga-Guide-Linear-Algebra/dp/1593274130/ref=as_li_ss_tl?ie=UTF8&qid=1452829566&sr=8-1&keywords=The+Manga+Guide+to+Linear+Algebra&linkCode=sl1&tag=theopesoudats-20&linkId=3d46150bdab2ef0ba9805cee31eb4d2d) — This is a fun filled Linear Algebra book which keeps Machine Learning in context. You will never forget these Algebra lessons for sure.

**Structured Thinking — 2 weeks**

* **Articles (mandatory):**These articles will guide you to structure your thinking process to approach problems in a better way so as to improve your efficiency.
* [How to train your mind for analytical thinking?](https://www.analyticsvidhya.com/blog/2014/01/train-mind-analytical-thinking/)
* [Tools for improving structured thinking](https://www.analyticsvidhya.com/blog/2014/02/tools-structured-thinking/)
* [The art of structured thinking and analyzing](https://www.analyticsvidhya.com/blog/2013/06/art-structured-thinking-analyzing/)

**Competitions (mandatory):**No amount of theory can beat practice. This is a [strategic thinking problem](https://datahack.analyticsvidhya.com/contest/practice-problem-strategic-thinking-ii/) which will test you on your thinking process. Also, keep an eye on business case studies as they help in structuring your thoughts tremendously.

**Step:3 Python Packages Pandas,numpy, matplotlib, scikit learn, bokeh**

This is where fun begins! Here is a brief introduction to various libraries. Let’s start practicing some common operations.

* Practice the [NumPy tutorial](http://wiki.scipy.org/Tentative_NumPy_Tutorial) thoroughly, especially NumPy arrays. This will form a good foundation for things to come.
* Next, look at the [SciPy tutorials](http://docs.scipy.org/doc/scipy/reference/tutorial/). Go through the introduction and the basics and do the remaining ones basis your needs.
* If you guessed Matplotlib tutorials next, you are wrong! They are too comprehensive for our need here. Instead look at this [ipython notebook](http://nbviewer.ipython.org/github/jrjohansson/scientific-python-lectures/blob/master/Lecture-4-Matplotlib.ipynb" \t "_blank) till Line 68 (i.e. till animations)
* Finally, let us look at Pandas. Pandas provide DataFrame functionality (like R) for Python. This is also where you should spend good time practicing. Pandas would become the most effective tool for all mid-size data analysis. Start with a short introduction, [10 minutes to pandas](http://pandas.pydata.org/pandas-docs/stable/10min.html). Then move on to a more detailed [tutorial on pandas](http://www.gregreda.com/2013/10/26/intro-to-pandas-data-structures/).
* Check out DataCamp’s course on [Pandas Foundations](https://www.datacamp.com/courses/pandas-foundations)

You can also look at [Exploratory Data Analysis with Pandas](https://www.analyticsvidhya.com/blog/2014/08/baby-steps-python-performing-exploratory-analysis-python/) and [Data munging with Pandas](https://www.analyticsvidhya.com/blog/2014/09/data-munging-python-using-pandas-baby-steps-python/)

**Additional Resources:**

* If you need a book on Pandas and NumPy, “[Python for Data Analysis](http://www.amazon.com/Python-Data-Analysis-Wrangling-IPython/dp/1449319793) by Wes McKinney”
* There are a lot of tutorials as part of Pandas documentation. You can have a look at them [here](http://pandas.pydata.org/pandas-docs/stable/tutorials.html)

**Assignment:**Solve this [assignment from CS109 course](http://nbviewer.ipython.org/github/cs109/2014/blob/master/homework/HW1.ipynb) from Harvard.

**Step : 4 Exploration and Visualization**

**1. R Programming**

**Course**

* [Exploratory Data Analysis](https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-learn-data-science-python-scratch-2/) — This is an awesome course by Johns Hopkins University on Coursera. You will need no other course to perform visualization and exploratory work in R.

**Blogs/Articles**

* [Comprehensive guide to Data Exploration in R](https://www.analyticsvidhya.com/blog/2015/04/comprehensive-guide-data-exploration-r/) — This will be a one-stop article that I will suggest you to go through carefully and follow every step. This is because the steps mentioned in the article are the same steps you will be using while solving any data problem or a hackathon problem.
* [Cheat sheet — Data Exploration in R](https://www.analyticsvidhya.com/blog/2015/10/cheatsheet-11-steps-data-exploration-with-codes/) — This cheat sheet contains all the steps in data exploration with codes. I suggest you to take out a print and paste it on your wall for quick reference.

**2. Python**

**Course (optional)**

* [Intro to Data Analysis](https://www.udacity.com/course/intro-to-data-analysis--ud170) — This is an excellent course by Udacity on Data Exploration using Numpy and Pandas.

**Blogs/Articles (mandatory)**

* [Comprehensive guide to Data Exploration using Python NumPy, Matplotlib and Pandas](https://www.analyticsvidhya.com/blog/2015/04/comprehensive-guide-data-exploration-sas-using-python-numpy-scipy-matplotlib-pandas/) — This is a sufficient and comprehensive article which uses the most popular Python libraries for exploration and visualization purposes.
* [9 popular ways to perform Data Visualization in Python](https://www.analyticsvidhya.com/blog/2015/05/data-visualization-python/) — This article presents the most commonly used graphs and plots used in Data Exploration along with Python codes. This is a must bookmarked article for people working in Data Science using Python.

**Books (optional)** — [Python for Data Analysis](http://www.cin.ufpe.br/~embat/Python%20for%20Data%20Analysis.pdf) — A one stop solution for your Data Exploration and Visualization in Python.

**Step : 5 Data Preprocessing**

1. [Data Preprocessing Story](https://medium.com/ml-research-lab/chapter-5-story-behind-data-preprocessing-799c06d8581d)
2. [Data Preprocessing Tutorial by Analytics Vidhya](https://www.analyticsvidhya.com/blog/2016/07/practical-guide-data-preprocessing-python-scikit-learn/)
3. [data Preprocessing](https://www.youtube.com/watch?v=NWp6DFtnqYk&list=PLjCAGuobkaDDFji_7NsgXMqxSPF30LgEw)

**Step:6 Feature Selection/ Engineering**

* **Course :**[**How to win the kaggle competition**](https://www.coursera.org/learn/competitive-data-science)
* **Blog** — [A Comprehensive Guide to Data Exploration](https://www.analyticsvidhya.com/blog/2016/01/guide-data-exploration/): This article will explain underlying techniques of feature engineering and different methods for feature creation
* **Books (optional)** **—**[Mastering Feature Engineering](https://www.amazon.com/Mastering-Feature-Engineering-Principles-Techniques/dp/1491953241): This book is master piece to learn feature engineering. Not only will you learn how to implement feature engineering in a systematic way. You will also learn different methods involved in feature engineering.

**Step: 7 Basic and Advance Machine learning algorithms**

**Basic Machine Learning Algorithms.**

* Linear Regression
* Logistic Regression
* Decision Trees
* KNN (K- Nearest Neighbours)
* K-Means
* Naïve Bayes
* Dimensionality Reduction

**Advanced algorithms**

* Random Forests
* Dimensionality Reduction Techniques
* Support Vector Machines
* Gradient Boosting Machines
* XGBOOST

**Linear Regression**

**Course**

* [Machine Learning by Andrew Ng](https://www.coursera.org/learn/machine-learning) — There is no better resource to learn Linear Regression than this course. It will give you a thorough understanding of linear regression and there is a reason why Andrew Ng is considered the rockstar of Machine Learning.

**Blogs/Articles**

* [This](https://onlinecourses.science.psu.edu/stat501/node/250) lesson out of PennState Stat 501 course outlines the main features of Linear Regression ranging from a simple definition of a Linear Regression to determining the goodness of fit of a regression line.
* [This](https://www.analyticsvidhya.com/blog/2015/10/regression-python-beginners/) is an excellent article with practical examples to explain Linear Regression with code.

**Books**

* [The Elements of Statistical Learning](http://www-stat.stanford.edu/~tibs/ElemStatLearn/printings/ESLII_print10.pdf) — This book is sometimes considered the holy grail of Machine Learning and Data Science. It explains Machine Learning concepts mathematically from a Statistics perspective.
* [Machine Learning with R](https://www.packtpub.com/big-data-and-business-intelligence/machine-learning-r) — This is a book I personally use to have a brief understanding of Machine Learning algorithms along with their implementation code.

**Practice**

* [Black Friday](https://datahack.analyticsvidhya.com/contest/black-friday/) — Like I already said — No amount of theory can beat practice. Here is a regression problem that you can try your hands on for a deeper understanding.

**Logistic Regression**

**Course (mandatory)**

* [Machine Learning by Andrew Ng](https://www.coursera.org/learn/machine-learning)– The week 3 of this course will give you a deeper understanding of the one of the most widely used classification algorithm.
* [Machine Learning: Classification](https://www.coursera.org/learn/ml-classification) — Week 1 and 2 of this practical oriented Specialization course using Python will satiate your knowledge thirst about Logistic Regression.

**Blogs/Articles (optional)**

* [Logistic Regression by Machine Learning Mastery](http://machinelearningmastery.com/logistic-regression-for-machine-learning/) — This is an excellent non-code based approach to Logistic regression to deepen your knowledge. I suggest you to have a look at it.

**Books (optional)**

* [Introduction to Statistical Learning](http://www-bcf.usc.edu/~gareth/ISL/) — This is an excellent book with a quality content on Logistic Regression’s underlying assumptions, statistical nature and mathematical linkage.

**Practice (mandatory)**

* [Loan Prediction](https://datahack.analyticsvidhya.com/contest/practice-problem-loan-prediction-iii/) — This is an excellent competition to practice and test your new Logistic Regression skills to predict whether loan status for a person was approved or not.

**Decision Trees**

**Course (mandatory)**

* [Machine Learning: Classification](https://www.coursera.org/learn/ml-classification) — Week 3 and 4 in this course is about the working of decision trees, preventing overfitting and handling missing values

**Blogs/Articles (mandatory)**

* [Technical Overview of decision trees](http://www.kdnuggets.com/2016/10/decision-trees-concise-technical-overview.html) — This is a quick overview of decision trees and a must read for anyone new to decision trees.
* [Complete tutorial on tree based modeling](https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-in-python/) — This is a python based tutorial on decision trees. For the sake of decision trees, read only sections 1–6 in this article.

**Books (mandatory)**

* [Introduction to Statistical Learning](http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Sixth%20Printing.pdf) — Section 8.1 and 8.3 explain the basics of decision trees through theory and practical examples.
* [Machine Learning with R](https://www.packtpub.com/big-data-and-business-intelligence/machine-learning-r) — Chapter 5 of this book provides you the best explanation of Machine Learning Algorithms available in the market. Here, the decision trees are explained in an extremely non-intimidating and easier style.

**Practice (mandatory)**

* [Loan Prediction](https://datahack.analyticsvidhya.com/contest/practice-problem-loan-prediction-iii/) — This is an excellent competition to practice and test your new Logistic Regression skills to predict whether loan status for a person was approved or not.

**KNN (K- Nearest Neighbors)**

**Course (mandatory)**

* [Machine Learning — Clustering and Retrieval](https://www.coursera.org/learn/ml-clustering-and-retrieval): Week 2 of this course progresses to k-nearest neighbors from 1-nearest neighbor and also describes the best ways to approximate the nearest neighbors. It explains all the concepts of KNN using python.

**Blogs/Articles (mandatory)**

* [Introduction to k-nearest neighbors: simplified](https://www.analyticsvidhya.com/blog/2014/10/introduction-k-neighbours-algorithm-clustering/) — This basic article describes when to use KNN, the ways in which k can be chosen and the way in which KNN algorithm works.
* [Learning KNN algorithm using R](https://www.analyticsvidhya.com/blog/2015/08/learning-concept-knn-algorithms-programming/) — This article is a comprehensive guide to learning KNN with hands-on codes for future references.

**K-Means**

**Course**

* [Machine Learning Course — Unsupervised Learning with K-means algorithm](https://www.coursera.org/learn/machine-learning): Week 8 of this discusses how to use course how K-means algorithm is used for handling unstructured data.

**Blog**

* [An Introduction to Clustering and different methods of clustering](https://www.analyticsvidhya.com/blog/2016/11/an-introduction-to-clustering-and-different-methods-of-clustering/): In this article, you will learn what is k-means clustering and the intricacies involved in that. It will give you a step by step approach how K-means algorithm works.

**Naive Bayes**

**Course**

* [Intro to Machine Learning](https://www.udacity.com/course/intro-to-machine-learning--ud120): Take this course to see Naive Bayes in action. In this course, Sebastian Thrun has explained Naive Bayes in Simple English.

**Blog / Article**

* [6 Easy Steps to Learn Naive Bayes Algorithm (with code in Python)](https://www.analyticsvidhya.com/blog/2015/09/naive-bayes-explained/) : This article will take you through Naive Bayes algorithm in detail. In this guide, you will learn how Naive Bayes algorithm works, applications and many more. It will also give you hands-on knowledge of building a model using Naive Bayes.
* [Naive Bayes for Machine Learning](http://machinelearningmastery.com/naive-bayes-for-machine-learning/) : This is one of the most comprehensive articles I have come across. Go through this article to have a complete understanding of why naive bayes algorithm is important for machine learning.

**Dimensionality Reduction**

**Course**

* [Machine Learning — Dimensionality Reduction:](https://www.coursera.org/learn/machine-learning) Week 8 of this course will walk you through dimensionality reduction and how Principal Components Analysis can be used for data compression of complex data.

**Blog / Article**

* [Beginners Guide To Learn Dimension Reduction Techniques:](https://www.analyticsvidhya.com/blog/2015/07/dimension-reduction-methods/)In this article, you will learn why dimension reduction is important in machine learning and the various techniques of dimension reduction.

**Random Forests**

* **Videos (mandatory)**
* [How Random Forest algorithm works?](https://www.youtube.com/watch?v=loNcrMjYh64) — Watch this video to have a visual perspective of how the Random Forest algorithm works.

**Books (optional)**

* [Introduction to Statistical Learning](http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Sixth%20Printing.pdf) — Section 8 explains the basics of Random Forests including bagging and boosting through theory and practical examples.
* [Applied predictive modeling](http://appliedpredictivemodeling.com/) — Chapter 8

**Blogs/Articles (mandatory)**

* [A tutorial on tree based modeling from scratch](https://www.analyticsvidhya.com/blog/2016/04/complete-tutorial-tree-based-modeling-scratch-in-python/) — This is an excellent article on trees based modeling using python. I suggest you to bookmark it right now.
* [Random Forests](https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm) — This blog explains the entire working, nuts and bolts of Random Forest.

**Gradient Boosting Machines**

**Blogs/Articles (mandatory)**

* [Guide on Boosting methods](https://www.analyticsvidhya.com/blog/2015/09/complete-guide-boosting-methods/)
* [Parameter tuning GBM](https://www.analyticsvidhya.com/blog/2016/02/complete-guide-parameter-tuning-gradient-boosting-gbm-python/)
* [Machine Learning Mastery- GBM](http://machinelearningmastery.com/gentle-introduction-gradient-boosting-algorithm-machine-learning/)

**Presentation (mandatory):**[Here](http://www.slideshare.net/mark_landry/gbm-package-in-r) is an excellent presentation on GBM. It contains the prominent features of GBM and the advantages and disadvantages of using it to solve real-world problems. It is must see article for somebody trying to understand GBM.

**XGBOOST**

**Blogs /Articles (mandatory)**

* [Official Introduction XGBOOST](http://xgboost.readthedocs.io/en/latest/model.html) — Read the documentation of hackathons winning algorithm. It is an improvement over GBM and is right now the most widely used algorithm for winning competitions.
* [Using XGBOOST in R](https://www.analyticsvidhya.com/blog/2016/01/xgboost-algorithm-easy-steps/) — An excellent article on deploying XGBOOST in R using a practical problem at hand.
* [XGBOOST for applied Machine Learning](http://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/) — An article by Machine Learning Mastery to evaluate the performance of XGBOOST over other algorithms.

**Support Vector Machines**

**Course (mandatory)**

* [Machine Learning by Andrew Ng](https://www.coursera.org/learn/machine-learning/home/week/7) — Week 7 of this course is an interesting place to start your SVM journey.

**Books (mandatory)**

* [Introduction to Statistical Learning](http://machinelearningmastery.com/gentle-introduction-xgboost-applied-machine-learning/)— Chapter 9 of the book contains a detail discussion about SVMs and the ways to deploy them.

**Blogs/Articles (optional)**

* [Understanding support vector machines](https://www.analyticsvidhya.com/blog/2015/10/understaing-support-vector-machine-example-code/) — This is an excellent article to understand an algorithm practically using examples.
* [SVM by Machine Learning Mastery](http://machinelearningmastery.com/support-vector-machines-for-machine-learning/) — This article discusses the different types of kernels employed in SVM and their uses.

**Step : 8 Profile Building on Github and Participation in Competition**

It is very important for a Data Scientist to have a GitHub profile to host all the codes of the project he/she has undertaken. Potential employers not only see what you have done, how you have coded and how frequently / how long you have been practicing data science.

Also, codes on GitHub open up avenues for open source projects which can highly boost your learning. If you don’t know how to use Git, you can learn from [Git and GitHub](https://www.udacity.com/course/how-to-use-git-and-github--ud775) on Udacity. This is one of the best and easy to learn course to manage the repositories through terminal.

1. [Analytics Vidhya Datahack](https://datahack.analyticsvidhya.com/contest/all/)
2. [Kaggle competitions](https://www.kaggle.com/competitions)
3. [Crowd Analytix human layer](https://www.crowdanalytix.com/community)
4. [DrivenData](https://www.drivendata.org/)

**Step: 9 Learn Some Advance Algorithm of Machine Learning**

There are a few specific machine learning algorithms, which come in handy while solving specific problems. For example, try solving online click prediction on large data sets with out applying online learning algorithms and you would know what I am talking about. Here are a few advanced ML algorithms you should learn this month:

**Online Machine Learning**

Course: [Online Methods In Machine Learning by MIT](http://www.mit.edu/~rakhlin/6.883/)

Books:

* [Understanding Machine Learning: From Theory to Algorithms](http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/index.html)
* [Online Learning and Online Convex Optimization](http://www.cs.huji.ac.il/~shais/papers/OLsurvey.pdf)

Blogs : [Langford’s hunch.net](http://hunch.net/)

**Vowpal Wabbit**

* Tutorial : [By the creator of Vowpal Wabbit — John Langford](https://github.com/JohnLangford/vowpal_wabbit/wiki/Tutorial)

**FTRL- Algorithms**

* Course: [Online Learning by University of Washington](http://courses.cs.washington.edu/courses/cse599s/12sp/index.html)
* Paper : [Google’s ad click prediction](http://www.eecs.tufts.edu/~dsculley/papers/ad-click-prediction.pdf)

**Exercise:** Practice on one of the old Kaggle competitions or open click through rate data sets as provided by Criteo.

**Step : 10 Deep Learning Basics & Advanced**

**Deep Learning Basics (May 2017 — June 2017)**

**Course (mandatory)**

* [Machine Learning by Andrew Ng](https://www.coursera.org/learn/machine-learning) — There is no better introductory material to Deep Learning and Neural Networks than Week 4 and Week 5 material of this course.
* [Deep learning by Google | Udacity](https://www.udacity.com/course/deep-learning--ud730) — This is an excellent basic course on transition from Machine Learning to Deep Learning, deep neural networks, Convolutional Neural Networks and Deep Learning for texts.

**Reading Material/Books**

* [Deep learning Textbook](http://www.deeplearningbook.org/) — Written by people like Ian Goodfellow, Yoshua Bengio and Aaron Courville, this book is bound to become the de-facto for people trying to learn Deep Learning.
* [Stanford Deep Learning tutorial](http://deeplearning.stanford.edu/tutorial/) — This is an all text and images resource provided by Stanford which starts from Linear Regression and goes to Convolutional Neural Networks with ease.

**Practice**— [Identify the digits](https://datahack.analyticsvidhya.com/contest/practice-problem-identify-the-digits/) — An awesome contest to check the basics you have learned to identify handwritten digits.

**Deep Learning advanced (June 2017 — August 2017)**

**Course (mandatory)**

* [Deep Learning by Oxford](https://www.cs.ox.ac.uk/people/nando.defreitas/machinelearning/)
* [Deep learning summer school at Montreal 2016](http://videolectures.net/deeplearning2016_montreal/) — This is a treasure trove of knowledge with many experts researching in the field of Deep Learning delivering keynote lectures.

**Specialization Material**

* **Deep Learning for Computer Vision**
* Primer: [“DL for Computer Vision”](http://www.analyticsvidhya.com/blog/2016/04/deep-learning-computer-vision-introduction-convolution-neural-networks/)
* Project: [“Facial Keypoint Detection”](http://danielnouri.org/notes/2014/12/17/using-convolutional-neural-nets-to-detect-facial-keypoints-tutorial/) Tutorial
* Required libraries: Nolearn
* Associated Course: [“CS231n: Convolutional Neural Networks for Visual Recognition”](http://cs231n.stanford.edu/)

**Deep Learning for Natural Language Processing**

* Primer : [“Deep Learning, NLP, and Representations”](http://colah.github.io/posts/2014-07-NLP-RNNs-Representations/)
* Project : “Deep Learning for Chatbots”: [“Part 1”](http://www.wildml.com/2016/04/deep-learning-for-chatbots-part-1-introduction/), [“Part 2”](http://www.wildml.com/2016/07/deep-learning-for-chatbots-2-retrieval-based-model-tensorflow/)
* Required library : Tensorflow
* Associated Course : [“CS224d: Deep Learning for Natural Language Processing”](http://cs224d.stanford.edu/)

**Deep Learning for Speech/Audio**

* Primer : [“Deep Speech: Lessons from Deep Learning”](http://usa.baidu.com/deep-speech-lessons-from-deep-learning/) news article and corresponding video.
* Project : [“Music Generation using Magenta (Tensorflow)”](https://magenta.tensorflow.org/2016/06/10/recurrent-neural-network-generation-tutorial/)
* Required library : Magenta
* Associated Course : [“Deep Learning (Spring 2016), CILVR Lab@NYU”](http://cilvr.cs.nyu.edu/doku.php?id=courses:deeplearning2016:start)

**Step 11 : Reinforcement Learning**

**Topics to be covered: Reinforcement Learning (Theory)**

**Course**

* [Course on Reinforcement Learning by David Silver UCL [Theory]](http://www0.cs.ucl.ac.uk/staff/d.silver/web/Teaching.html)
* [Udacity — Machine learning Reinforcement Learning](https://www.udacity.com/course/machine-learning-reinforcement-learning--ud820) [Theory +Example]
* [Reinforcement Learning video by Andrew NG](https://www.youtube.com/watch?v=RtxI449ZjSc&feature=relmfu) [Theory]

**Code**[Reinforcement Learning Introductory Codes](https://github.com/ShangtongZhang/reinforcement-learning-an-introduction)[Code]

* **Books**[Reinforcement Learning by MIT press](https://mitpress.mit.edu/books/reinforcement-learning) — This will be good reference material for the reinforcement learning taught by the professors at MIT.

**Competitions:**

* [An awesome website for reinforcement Learning Competitions](http://www.rl-competition.org/)
* [Kaggle’s first reinforcement style learning competition- Two sigma](https://www.kaggle.com/c/two-sigma-financial-modeling)

**References :**

1. [https://www.analyticsvidhya.com/blog/2017/01/the-most-comprehensive-data-science-learning-plan-for-2017](https://www.analyticsvidhya.com/blog/2017/01/the-most-comprehensive-data-science-learning-plan-for-2017/#five-six)/
2. <https://www.analyticsvidhya.com/learning-paths-data-science-business-analytics-business-intelligence-big-data/learning-path-data-science-python/>
3. <https://dzone.com/articles/10-steps-to-become-data-scientist-in-2018>
4. <https://medium.freecodecamp.org/a-path-for-you-to-learn-analytics-and-data-skills-bd48ccde7325>
5. <https://www.quora.com/How-can-I-become-a-data-scientist-1>