THE DATA SCIENTIST ROADMAP

Math fundamentals

You need to learn the concepts of Linear algebra, calculus, optimisation and functions. You don't need a PHD just knowing the basics is enough.

Resources

- Part 1 of deeplearningbook.org
- Math track of Khan academy
- mml-book.com
- Linear algebra by fast.ai
- For deeper understanding go with MIT OpenCourseware

Learn programming

Choose any programming language (Go with Python) and learn interactively. Also, learn CS fundamentals.

Resources

- CS50x by harvard
- learnpython.org
- Python crash course
- Learn python the hard way
- pythonprogramming.net/

Data wrangling and visualization

Understanding Data wrangling techniques (data collection, cleaning and exploration) is very important as a data scientist. Also, learn how to create and study the visual representation of the data.

Resources

- Python for Data Analysis by Wes McKinney
- Few amazing blongs on medium and analyticsvidhya
- Getting and cleaning data by John Hopkins (Coursera)
- Hands-On Data Analysis with Pandas

statistics and probability

learn descriptive and inferential, associative, and differential statistical concepts and also probability (Conditional Probability and Bayes' Theorem).

Resources

- Khan academy
- Intro to statistics by udacity
- MIT OpenCourseware
- Introduction to statistical learning or elements of statistical learning

Understand Databases

Most of the companies use relational databases instead of text files to store data. It can be any like MySQL, Postgres, MongoDB etc...

Resources

- MongoDB university
- Getting Started with SQL: A Hands-On Approach for Beginners
- SQL for Data Science (Coursera)
- SQL and Database course by freeCodeCamp
- etc....

Learn machine learning

lo start Machine Learning, first understand the terminologies around machine learning and its types (supervised learning, unsupervised learning, Dimensionality reduction techniques, time series etc....

Resources

- Machine Learning by Andrew Ng (Coursera)
- Fast.ai Machine learning course
- Practical machine learning by John Hopkins (Coursera)
- mlcourse.ai
- Hands-On Machine Learning with Scikit-Learn and TensorFlow

Practice and practice

Get good hands-on experience by practising what you learned by doing projects, participate in competitions, meet fellow data scientists and learn from them.

Resources

- Kaggle, Analyticsvidhya, Datadriven, Crowd AI etc...
- Fast.ai forum, meetup, our own community, hackathons
- Linkedin and github
- etc...

Learn Big Data

A Single machine is not enough when you are working with huge amounts of data. Learn Big data fundamentals, Hadoop Ecosystem, MapReduce, Apache Spark etc....Also, learn how to deploy the models and maintain them.

Resources

- Big Data Specialization by UCSanDiego (Coursera)
- Getting started with Apache Spark by James A. Scott
- Respective documentations/websites.
- fullstackdeeplearning.com

Get a job

Best way to test your skills is to work on a real-life problem. You can get a job/internship or take a bootcamp or work on your own venture where you can do this. Start with an internship, Bootcamp and get a full-time.

Resources

- Create a perfect resume before applying to the jobs.
- Find jobs here (linkedin, angel.co, stackoverflow, kaggle etc...)
- Take any BootCamps (I suggest to join Dphi.tech)
- Follow and engage with the community (share your knowledge)

Learn advance concepts

It's time to learn advance concepts (Deep learning). Based on your job requirements or your own project, learn the related concepts.

Resources

- Deeplearning.ai
- Part 2 of deeplearningbook.org
- TensorFlow or PyTorch
- fast.ai