MITx: 6.86x Machine Learning with Python - From Linear Models to Deep Learning

Overview

December 2019

Comment: This overview has been written in R Markdown that provides an authoring framework for data science. Learn more at https://bookdown.org/yihui/rmarkdown/

MITx's MicroMasters Program in Statistics and Data Science

From probability and statistics to data analysis and machine learning, master the skills needed to solve complex challenges with data.

In september 2019. a great MIT's movement, to open their doors to the whole world, has come to the field of Data Science. Anyone from the EdX platform for online learning, showing interests for Data Science has gotten an invitation to join. In the next year or so, the future students were asked to pass 4 MIT exams and one Capstone exam, so they could get the MIT's credentials in Data Science area. If they achieve the goal, they could also apply for the second semester on MIT's campus and finish **Masters of Statistics and Data Science** in a hybrid generation of online and on campus students.

Mastering the foundations of data science, statistics, and machine learning using programming languages like **Python** and **R** students are prepared for jobs like: Data Scientist, Data Analyst, Business Intelligence Analyst, Systems Analyst, Data Engineer. More about the program could be found **here**.

MITx: 6.86x Machine Learning with Python

An in-depth introduction to the field of machine learning, from linear models to deep learning and reinforcement learning, through hands-on Python projects. – Course 4 of 4 in the MITx MicroMasters program in Statistics and Data Science.

MITx: 6.86x is advanced 14 week long instructor paced online course given by MIT's Department of Electrical Engineering and Computer Science. Prerequisites for this course are:

- MIT: 6.00.1x (Introduction to Computer Science and Programming Using Python) or proficiency in Python programming
- MITx: 6.431x (Probability - The Science of Uncertainty and Data) or equivalent probability theory course
- College-level single and multi-variable calculus
- Vectors and matrices

Course is presented in 6 Units, with 19 Lectures, 7 Homeworks and 5 Python Projects, through topics like:

- Representation, over-fitting, regularization, generalization, VC dimension
- Clustering, classification, recommender problems, probabilistic modeling, reinforcement learning
- On-line algorithms, support vector machines, and neural networks/deep learning

where stidents are gaining knowledge about principles and algorithms for turning training data into effective automated predictions. In it's second edition, the course already has 50 thousands students enrolled.

Unit 0: Course overview

Brief Review of Vectors, Planes, and Optimization

In this Unit, three videos (total 21:09 minutes) were presented showing basic overview of:

- **vectors** in n-dimensional space (addition and subtraction of two vectors, norm of an vector, dot product of two vectors)
- **plane definition** in n-dimensional spaces (used later in classification for decision boundary) as: a point x of n-dimensional space belongs to some plane if it satisfies the equation $\theta \cdot x + \theta_0 = 0$, where θ is a vector perpendicular to that plane and θ_0 an offset of the plane from it's origin.
- loss Function, gradient descent and chain rule