Optimization For Data Science

Description

This course introduces the students to optimization problems and help them understand the different algorithms the modern data science libraries use, especially the gradient descent, its variants and some momentum-based methods. The idea is to have a better grasp of the concepts which make data science tools so efficient without using them as black boxes.

Learning Objectives and Outcomes

- Understand convexity
- Understand optimization problems
- Understand the different variants of the gradient descent algorithm
- Understand the momentum based algorithms
- Be able to code an implementation of the gradient descent algorithm

Course Schedule and Contents

Session #1

- Reminder about data science and machine learning
- Definition of optimization problems
  → understand what an optimization problem is
  → understand what a local/global extremum is
- Definition of convexity
  → understand what are convex sets, convex functions and convex problems
  → understand why convexity is useful in optimization

Session #2

- Definition of supervised learning
  → understand supervised learning, training sets
- Definition of linear regression
  → understand what a model is, the objective function
- Reminder about the gradient
- Intuition behind the gradient descent
  → understand the algorithm without the mathematical aspects
- Explanation of the Gradient descent algorithm
  → understand the learning rate and how the gradient is used to minimize the cost function

Optimization For Data Science

Session #3	▪ Learning rate issues → understand the vanishing gradient, values too big/small for alpha ▪ Variants of Gradient descent → understand the vanilla GD, stochastic GD and mini batch GD ▪ Accelerating the GD (part I) → understand the momentum algorithm, NAG, Adagrad, Adam, RMSProp
Session #4	▪ Accelerating the GD (part II) → understand the momentum algorithm, NAG, Adagrad, Adam, RMSProp ▪ Logistic regression and classification → understand the classification problems and the use of the logistic function ▪ Neural networks → understand the connection between the NN, the sigmoid function and the gradient descent algorithm ▪ Newton method → understand this method to find the zeros of a function and the connection with descent methods ▪ Unsupervised learning → Understand clustering ▪ Dimensionality reduction → understand the PCA ▪ Project → Assignment for the next session
Session #5	▪ Project review

→ being able to explain the project and to answer a few questions about it

Grading

Course Project: 100%

Policies

- I expect you to submit your reports on time to receive proper credit/grade.
- Any work submitted must be your own.
- I expect everyone to contribute equally to group assignments
- Attendance in every class is expected. Class participation and discussion are strongly encouraged.
- Late work will not be accepted unless prior arrangements have been made directly with me.
- Cases will be decided on an individual basis.

Good Luck!