Data Analysis Project (DAP)

Overview

One of the main goals of the course is to prepare you to apply machine learning algorithms to real-world tasks on finding interesting patterns or providing predictions, and to leave you well-qualified to conduct machine learning research or build applications. This project is intended to start you in these directions.

As the course progresses you will be building a data analysis project for a real-world scenario of your choosing. You will pick a project topic, a machine learning application that interests you, define problems for which you plan to apply machine learning techniques, and identify one or several datasets suitable for your topic of interest. You will then explore how best to apply machine learning algorithms to solve it: preprocess the data, conduct an exploratory data analysis, develop models using supervised learning and unsupervised learning algorithms, and refine the models for performance improvement.

The project consists of seven sequencing assignments. In order to move on the next step of the project, you have to complete the previous assignments. The brief description and due dates of the seven assignments are as follows:

(1) DAP 1: **Project design and milestone (50 pts)**
- Choose a project topic of your interest and design a milestone of the project
- Due: Friday, April 17th

(2) DAP 2: **Data preprocessing and exploratory data analysis (50 pts)**
- Preprocess the data and conduct exploratory data analysis.
- Due: Friday, April 24th

(3) DAP 3: Classification (50 pts)
- Apply commonly used classification algorithms and evaluate models.
- Due: Friday, May 1st

(4) DAP 4: Regression (50 pts)
- Develop regression models: linear regression and logistic regression.
- Due: Monday, May 11th

(5) DAP 5: Cluster analysis (50 pts)
- Conduct cluster analysis
- Due: Monday, May 18th

(6) DAP 6: Neural networks (50 pts)
- Design and implement a neural network.
- Due: Monday, June 1st

(7) Model refinement and data analysis report (100 pts)
- Due: Friday, June 5th

If you are looking for project ideas or if you are unsure whether your proposed application is appropriate, please come to my office hours, and I would be happy to brainstorm and suggest some project ideas.