CS 475/675 Machine Learning: Homework 3

Due: Friday, October 18, 2024, 11:59 pm US/Eastern

100 Points Total Version 1.0

Make sure to read from start to finish before beginning the assignment.

1 Homeworks

Homeworks will typically contain two parts:

1. **Analytical:** These analytical questions will consider topics from the course. These will include mathematical derivations and analyses. Your answers will be entirely based on written work, i.e. no programming.
2. **Practicum:** In the practicum portion of the assignment, you will apply machine learning concepts to gain experience working with data from different domains. Practicums could involve Python notebooks, applied explorations of topics covered in the class, or programming assignments. Please note that the use of any form of AI assistance is strictly prohibited in this assignment.

   Click here for the Practicum Google Colab Notebook

The point total for each portion of the homework will be listed in the assignment. Written assignments will be submitted as PDFs. See below for more details about what to submit.

1.1 Collaboration Policy

The course policy is that, *unless otherwise specified*, all work must be your own.

1.2 What to Submit

For this assignment you will submit the following.

1. **Analytical**. You will submit your analytical solutions to Gradescope. **Your writeup must be compiled from LaTeX and uploaded as a PDF**. The writeup should contain all of the answers to the analytical questions asked in the assignment. Make sure to include your name in the writeup PDF and to use the provided LaTeX template for your answers following the distributed template. You will submit this to the assignment called "Homework 3: Analytical".
2. **Practicum Python Notebook.** You will submit the notebook (*.ipynb) to the assignment titled "Homework 3: Practicum".

You will need to create an account on gradescope.com and signup for this class. The course is `https://www.gradescope.com/courses/835426`. Use entry code `BK5K8K`. **You must either use the email account associated with your JHED, or specify your JHED as your student ID.** See this video for instructions on how to upload a homework assignment: `https://www.youtube.com/watch?v=KMPoby5g_nE`.

1.3 Questions?

Remember to submit questions about the assignment to courselore: `courselore.org/courses/9582956601/`.

2 Analytical (50 points)

Please see the accompanying `2024_homework3_analytical.tex` file for the analytical questions for this assignment. There is space provided in that file for you to type your answers in LaTeXafter each question. **Do not edit the file in any way except to add your answers.** Gradescope assumes that the PDF will exactly match our template except for your solutions.

3 Practicum (50 points)

In this assignment, we will implement **Perceptron**, and **Support Vector Machine**. The notebook is structured into two parts:

Part 1: Implement Perceptron using Rosenblatt's Algorithm

The perceptron, developed by Frank Rosenblatt in 1958, is a fundamental model in artificial neural networks that acts as a binary classifier. Implementing the perceptron using Rosenblatt's algorithm involves adjusting weights based on input data through a simple supervised learning rule. This iterative process minimizes the error between predicted and actual outputs, allowing the perceptron to learn from training examples. By understanding this algorithm, one gains insight into the foundational principles of machine learning and how basic neural networks operate to classify data effectively.

Part 2: Implement Support Vector Machine using Quadratic Solver

We will implement the Support Vector Machine (SVM) algorithm using the Quadratic Solver. The SVM algorithm works by finding the hyperplane that best separates the data points into different classes. As we've seen in class, we could turn the primal optimization problem into a dual optimization problem. Your task is to map this dual optimization problem into a Quadratic Programming problem and solve it using the `cvxopt` solver.

What You Will Do

Open the Jupyter notebook `CS475_homework3_practicum.ipynb`. This notebook will walk you through:

- Implementing the Perceptron algorithm using Rosenblatt's algorithm for binary classification task.
- Implementing the Support Vector Machine algorithm using the Quadratic Solver. The notebook will guide you through the process of mapping the dual optimization problem into a Quadratic Programming problem and solving it using the `cvxopt` solver. Understanding how kernels can be used to classify non-linearly separable data.

There are questions that should be answered inline within the notebook. You will hand in the Python notebook, which contains your implementation and the answers to the questions. Please see the instructions on how to submit the notebook.

Click here for the Practicum Google Colab Notebook