Home Assignment #2

Machine Learning and Deep Learning (CDSCO2041C)

Somnath Mazumdar
sma.digi@cbs.dk
Department of Digitalization, Copenhagen Business School

Deadline: 19-03-2025

Instructions

1. Write your complete name and student ID on the report.
2. The project report should be in pdf format of a maximum of 5 pages. The report should conform to the general formatting guidelines and academic standards that are expected for written projects at CBS.
3. If you have more content to present, feel free to include them in the Appendix to the report. But, not the complete code.
4. Complete solution code should be submitted as **one single *jupyter* notebook** and attach separately.

Question 1 : **Investment Strategy [Coding]**

Imagine you have DKK 500,000 saved in your bank account and you're considering investing in either cryptocurrency or stocks. You've chosen Ethereum over Bitcoin due to its cost advantage.

You're unsure which individual stocks to pick, but you have data on the S&P 500 (Standard & Poor's 500), a stock market index that tracks the performance of 500 of the largest publicly traded companies in the United States. You've opted for this data instead of individual stock data because the S&P 500 is commonly used as a benchmark to measure the performance of individual stocks and investment portfolios.

Now, based on Return on Investment (RoI)[a] metric, you want to decide:

1. Where to invest (e.g., cryptocurrency or stock)?
2. How much to invest?
3. When to invest? Is there any seasonality for cryptocurrency and stock? For instance, should you wait for a specific month before starting to invest?
4. Whether to use a pure or hybrid strategy (means split investment approach X% in cryptocurrency or Y% in stock)?
5. What should be the investment horizon (how long: months or years)?

For the analysis, use the data file "ethereum.csv" and "spx.csv" and your data range should be Feb 28, 2022 and Feb 28, 2025 (which is a three years). Hints: *Consider, 1 USD = DKK 7.10 and use the techniques/models which are covered till Lecture 8.*

[a] RoI is a financial metric used to evaluate the profitability of an investment. It measures how much return (profit) you earn relative to the cost of the investment.

Question 2 : **Gradient Descent [Pen and paper-based solution]**

1. What is a learner said to do when it outputs a classifier that is 100% accurate on the training data but only 50% accurate on test data, when in fact it could have output one that is 75% accurate on both?
2. In stochastic gradient descent, each pass over the dataset requires the same number of arithmetic operations, whether we use mini-batches of size one or size 1000. Why can it nevertheless be more computationally efficient to use mini-batches of size 1000?
3. Below Figure 1 shows the level curves in the weight space of a cost function C which we are trying to minimize. The current weight vector is marked by an x. Sketch the gradient descent update.

> **Hint:** We haven't given you enough information to determine the magnitude, so we want you to correct the direction.

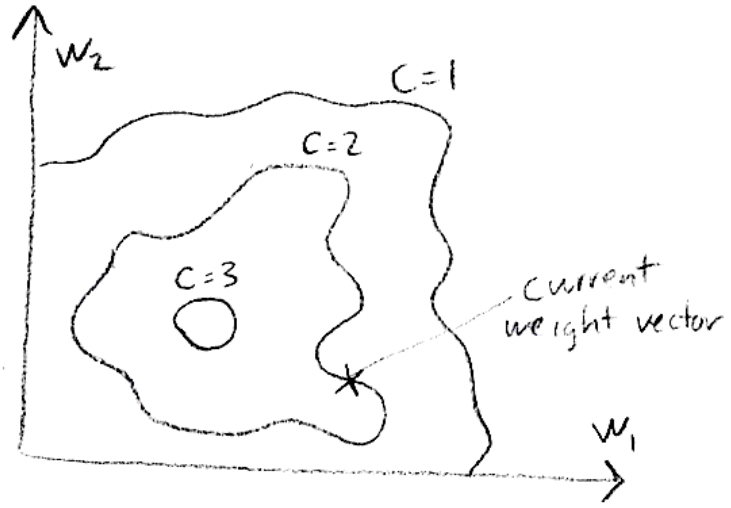

Figure 1: Use this diagram to sketch the gradient descent update