MATHEMATICS FOR MACHINE LEARNING
LAB 8 – 5%
Variance and Correlation

Solve the following questions. You can use Python or solve them on paper (Or you can solve some using Python and some on paper). If you are using Python to solve, please make sure to attach the output in your submissions. Like the other labs, please hand in PDF and/or Python code with output.

Question 1: A teacher wants to analyze the performance of students in a math test. Given the scores:

scores = [78, 85, 90, 92, 88, 76, 95]

Compute and print the variance to understand the spread of scores.

Question 2: The teacher also wants to measure how much the scores deviate from the average.

Compute and print the standard deviation of the scores.

Question 3: A company wants to analyze the relationship between employee experience (in years) and their productivity score.

experience = [1, 2, 3, 4, 5]

productivity = [50, 55, 65, 70, 80]

Compute and print Pearson correlation to assess the linear relationship.

Question 4: The HR department is interested in determining if there is a monotonic relationship between employee experience and productivity.

Compute and print Spearman correlation.

Question 5: A business analyst wants to visualize the relationship between advertising spending and sales.

ad_spending = [100, 200, 300, 400, 500]

sales = [20, 40, 60, 80, 100]

Generate and display a scatter plot.

Question 6: A health researcher is analyzing the correlation between daily exercise (minutes) and cholesterol levels.

df = pd.DataFrame({'Exercise': [30, 40, 50, 60, 70], 'Cholesterol': [220, 210, 200, 190, 180], 'Weight': [80, 78, 75, 73, 70]})

Generate and display a correlation heatmap to identify relationships.

Question 7: A financial analyst wants to analyze the volatility of stock prices over time.

stocks = [100, 102, 105, 98, 97, 110, 115]

Compute the rolling variance using a window size of 3.

Question 8: A data scientist wants to study the impact of study hours on exam scores and visualize the trend.

study_hours = [1, 2, 3, 4, 5]

exam_scores = [50, 60, 70, 80, 90]

Generate and display a scatter plot with a regression line.