Spring 2025 Total points: 100

COMP/EECE 7720/8720 Artificial Intelligence

Project Assignment

Due: Wednesday 4/30/2025 by 11:00 am.

Submission instruction: Do *not* email your submission to me. Upload your submission (code files and PDF report) as a single ZIP file in the Canvas folder "Project".

Name:_____ U-number:_____

Please write legibly and show all steps. State all assumptions clearly.

Stock price prediction problem: Given the stock price for the last *n* trading days, predict the stock price for the next trading day.

We will use the daily adjusted closing price "Adj Close" of the S&P 500 as the stock price. This data can be downloaded from Yahoo Finance: https://finance.yahoo.com/quote/%5EGSPC/history/ . We will use this data from December 30, 1927 through March 31, 2025.

We will use three algorithms to solve this problem: Hidden Markov Model (HMM), Kalman filter, and Dynamic Bayesian Network (DBN). Sample code along with explanation for these three algorithms is in the links below. You may or may not use this code.
HMM: https://www.kaggle.com/code/ehsanamim/stock-market-prediction-using-hmm
Kalman filter: https://medium.com/intro-to-artificial-intelligence/kalman-filter-in-stock-trading-552e1e4b2dfb
DBN: https://www.linkedin.com/pulse/bayesian-probability-stock-market-prediction-in-depth-anand-damdiyal-98nec/

In the report, include a plot of the prediction accuracy of the three algorithms for the following values of *n*: 5, 10, 20, 30, 50, 100, 200, 300, 500, 1000. That is, x-axis will be *n* and y-axis will be Average_MAPE (see definition below).

The prediction accuracy for a day should be calculated in terms of Mean Absolute Percentage Error (MAPE).

$$\text{MAPE}(k) = 100 \text{ x } |\text{Actual\_stock\_price}(k) - \text{Predicted\_stock\_price}(k)| \text{ / Actual\_stock\_price}(k)$$

where $|.|$ denotes absolute value, $\text{Actual\_stock\_price}(k)$ is the actual stock price on the k-th trading day, $\text{Predicted\_stock\_price}(k)$ is the predicted stock price on the k-th trading day, and $\text{MAPE}(k)$ is the MAPE for the k-th trading day. For prediction, shift the window of length *n* by one trading day at a time.

The prediction accuracy over the entire duration (12/30/1927—03/31/2025) should be calculated in terms of Average_MAPE.

$$\text{Average\_MAPE} = \frac{1}{N-n}\sum_{k=n+1}^{N} \text{MAPE}(k)$$

where N is the total number of trading days in the entire duration. The first prediction will be for the (n+1)-th trading day.

Use a widely-used language for coding. Python and Matlab would be more useful. Submit all code so that your reported results can be reproduced at my end. No points will be awarded if code fails to run on my end. This is an individual project. Strictly follow the submission procedure stated above. Late submissions will receive zero points (no exception).

Report 25 points + Code 75 points = 100 points.