COMP5313 Artificial Intelligence Department of Computer Science

Project 1: Stock Price Prediction Using AI and Machine Learning

The project is about predicting stock prices using machine learning and deep learning. I achieved the results using three models: firstly, I used the Arima model; secondly, I used the Sarimax. Model, and lastly, I used the Short-Term Memory model. I mentioned the comparisons of all the models below.

ARIMA Model:

The ARIMA (Autoregressive Integrated Moving Average) model is a time series forecasting approach that combines autoregressive (AR), differencing (I), and moving average (MA) components to model and predict future observations. The autoregressive component captures the linear relationship between the current observation and its past values while differencing transforms the time series to achieve stationarity. The moving average component models the relationship between the current observation and a residual error from a moving average model applied to lagged observations. The model is denoted as ARIMA (p, d, q), where 'p' is the order of autoregressive terms, 'd' is the order of differencing, and 'q' is the order of moving average terms. ARIMA is widely used for forecasting univariate time series data, although its effectiveness may vary depending on the underlying patterns in the data. Identifying appropriate orders, parameter estimation, model validation, and subsequent forecasting are key to utilizing ARIMA for time series analysis.

SARIMAX Model:

SARIMAX, which stands for Seasonal Autoregressive Integrated Moving Average with exogenous regressors, is an extension of the ARIMA model designed to handle time series data with seasonal patterns and external variables. SARIMAX incorporates additional components, allowing it to capture the influence of external factors on the time series. Like ARIMA, SARIMAX includes autoregressive (AR), differencing (I), and moving average (MA) components, but it also introduces seasonal components represented by the terms SAR (Seasonal Autoregressive) and SMA (Seasonal Moving Average). The 'X' in SARIMAX signifies the inclusion of exogenous variables, which are external factors that may impact the time series. The model is denoted as SARIMAX (p, d, q) (P, D, Q, s), where the lowercase letters represent non-seasonal orders, and the uppercase letters represent seasonal orders, with 's' denoting the length of the seasonal cycle. SARIMAX is particularly useful for forecasting time series data that exhibits both trend and seasonal patterns while considering the influence of external factors.

Long Short-Term Memory Model:

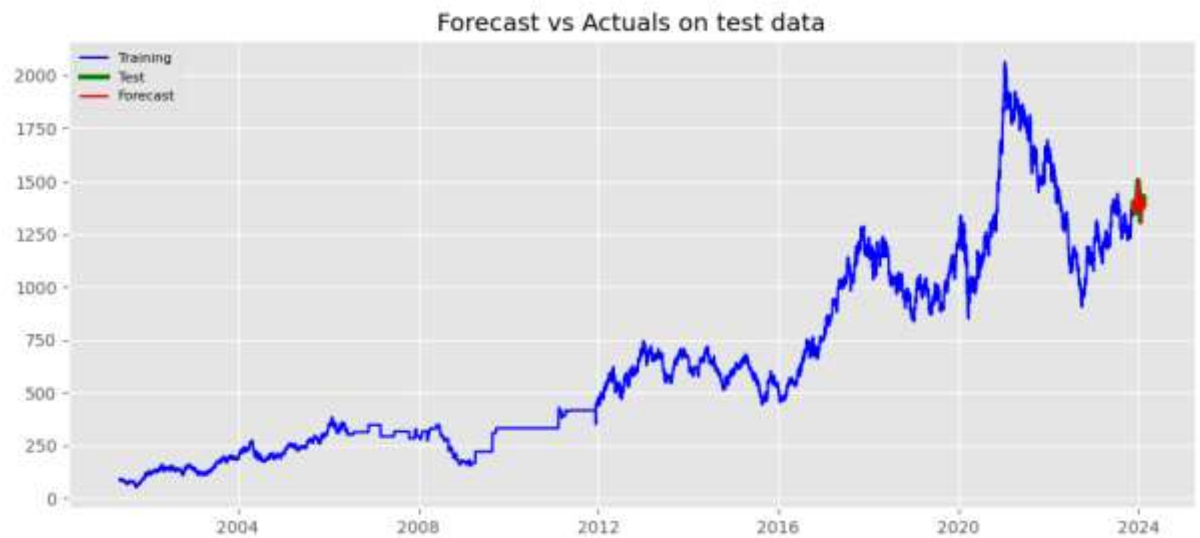
In the realm of time series problems, Long Short-Term Memory (LSTM) serves as a powerful neural network architecture specifically tailored to address the challenges associated with modeling and predicting sequential data over time. Time series data often involves intricate temporal dependencies and patterns, and LSTMs excel in capturing and learning these nuances. Unlike traditional recurrent neural networks (RNNs), LSTMs mitigate issues such as vanishing gradients by incorporating memory cells and gating mechanisms. This design enables LSTMs to selectively retain or discard information at each time step, preserving crucial long-term dependencies. Applied to time series forecasting, LSTMs can analyze historical sequences, discern patterns, and make accurate predictions for future timestamps, making them invaluable for tasks like stock price prediction, energy consumption forecasting, and other domains where understanding and leveraging temporal dynamics are paramount.

Comparisons:

Root Mean Squared Error:

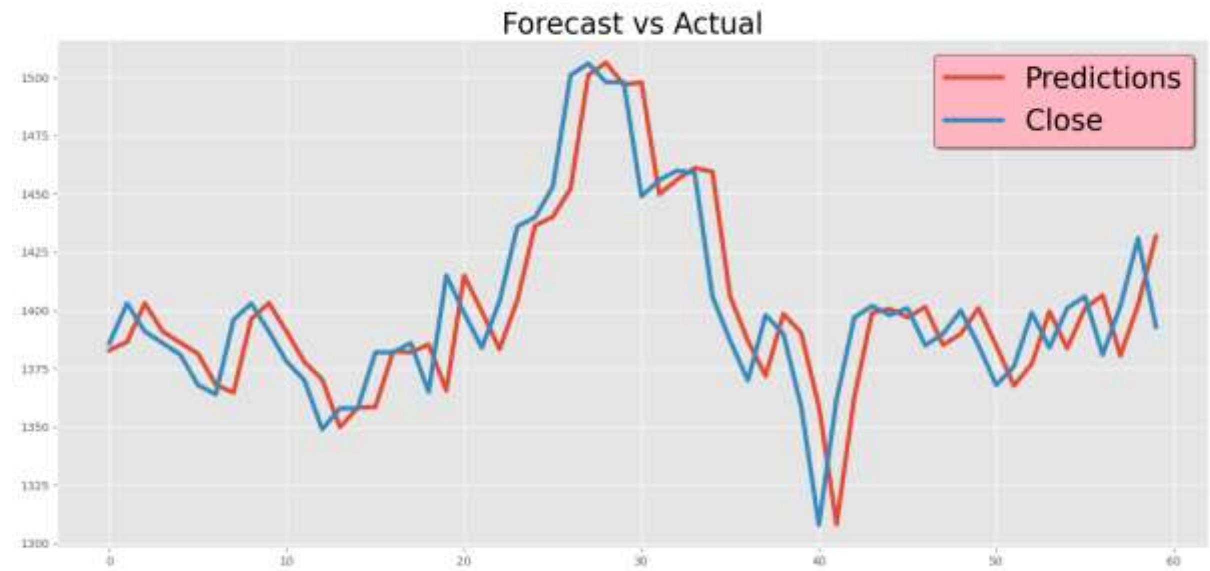
ARIMA	RMSE: 22.467
SARIMA	RMSE: 22.53100017792959
Long Short-Term Memory	RMSE: 0.021402547135949135

ARIMA:



ARIMA Model Predicting for the next 60 days:

SARIMA:



Long Short-Term Memory: