An Empirical Study on: Time Series Forecasting of Amazon Stock Prices using Neural Networks LSTM and GAN models

Anjul Bhardwaj
Dept. Artificial Intelligence & Data Science, Poornima Institute of Engineering & Technology, Jaipur, Rajasthan
anjulbhardwaj20july@gmail.com

Uday Pratap Singh
Dept. Artificial Intelligence & Data Science, Poornima Institute of Engineering & Technology, Jaipur, Rajasthan
udaykbcaos@gmail.com

ABSTRACT

The OHLCV (Open, High, Low, Close, Volume) data used in this study is used to forecast time series and anticipate stock price movement. We investigate a wide variety of models, including traditional statistical approaches and cutting-edge deep learning strategies combined with sentiment analysis, feature extraction, and hyperparameter tweaking. Instead of focusing on absolute stock prices, our main goal is to predict swings in stock prices, as this has been shown to produce more accurate outcomes. We start our research by obtaining historical Amazon stock data via the Yahoo API, and then we go on a thorough analytical journey. We generate features on the OHLCV data first, and then we design and test Fourier and Autoregressive Integrated Moving Average (ARIMA) models. We then switch to more sophisticated deep learning methods, using the pre-processed data to apply Long Short-Term Memory (LSTM) models. Interestingly, we add sentiment analysis to the LSTM study, which expands its scope and lets us consider market sentiment as a possible influencing factor. To guarantee the stability of our models, we use a careful train-test split technique and organize the data in a time series manner. The field of financial forecasting and trading methods will ultimately benefit from the insightful information our study's findings provide on the efficacy of different modeling techniques and their capacity to forecast stock price movements.

KEYWORDS

Financial Forecasting, Time series, Neural Networks, OHLV (Open, High, Low, Volume), LSTM (Long Short-Term Memory), ARIMA (Autoregressive Integrated Moving Average)

ACM Reference Format:
Anjul Bhardwaj and Uday Pratap Singh. 2023. An Empirical Study on: Time Series Forecasting of Amazon Stock Prices using Neural Networks LSTM and GAN models. In *International Conference on Information Management & Machine Intelligence (ICIMMI 2023), November 23–25, 2023, Jaipur, India.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3647444.3652440

1 INTRODUCTION

Forecasting stock prices with a high degree of accuracy has been a long-standing issue for investors, traders, and scholars alike because the financial markets are known for their inherent volatility and complexity. Advanced machine learning approaches are becoming more and more common in the field of financial forecasting in this era of data-driven decision-making. Using Long Short-Term Memory (LSTM) and Generative Adversarial Networks (GANs) in particular, this research paper presents an empirical investigation that aims to use the potential of neural networks for time series forecasting of Amazon stock prices. One of the biggest and most significant technological and e-commerce corporations in the world, Amazon, makes a great analytical subject.

It is a good choice for examining the efficacy of sophisticated forecasting models because of the wide range of factors that affect its stock prices, such as macroeconomic conditions, corporate performance, market sentiment, and world events.

This study's main goal is to investigate how well LSTM and GAN models can forecast the price of Amazon shares. For time series forecasting tasks, LSTM, a specialized recurrent neural network design, has gained widespread adoption due to its exceptional ability to capture sequential patterns. Conversely, GANs are well-known for their generating and discriminative powers, and it would be interesting to investigate their potential for simulating the dynamics of financial time series data. This study uses a thorough analytic strategy that involves sentiment analysis and feature engineering in addition to neural network models. An integrated comprehension of the dynamics operating in the financial markets is made possible by the incorporation of sentiment analysis as a potential driving force behind changes in stock prices.

To assure the stability and dependability of the models, the study methodology entails gathering historical Amazon stock data using the Yahoo API, pre-processing the data, and using a strict train-test split technique. Through a time, series analysis of this dataset, the research aims to reveal insights that will benefit practitioners and scholars working in the financial forecasting and trading fields. The results of this study should clarify how well LSTM and GAN models predict Amazon stock prices and offer insightful information about the field of financial forecasting. Developing precise and dependable forecasting techniques is essential for making well-informed decisions and investment plans as the financial markets keep changing. This paper represents a step towards harnessing the potential of cutting-edge machine learning models for addressing this challenge.

2 RELATED WORK

The fundamental analysis of a company, including its financial situation, quarterly balance, dividends, audit reports, sales data, import/export volume, and audit reports, is a part of stock market forecasting. This is just half of the picture; the other part consists

ICIMMI 2023, November 23–25, 2023, Jaipur, India

ACM ISBN 979-8-4007-0941-8/23/11
https://doi.org/10.1145/3647444.3652440

of opinions and rumours about the company in the actual world, which are crucial because they might affect stock price changes. The hybrid approach based on a model integrates all market-influencing factors. To make use of the benefits of each model in predicting weekly stock values, [1] create a Proposed Hybrid Model (PHM) that combines three models: the Exponential Smoothing Method (ESM), Autoregressive Integrated Moving Average (ARIMA), and a Backpropagation Neural Network (BPNN) model. This study shows that the hybrid model outperforms all component sub-models and conventional models with a directional accuracy of 70.16% when tested on the Shenzhen Integrated Index and Dow Jones Industrial Average (DJIA). A hybrid deep learning model that combines the well-known Deep Neural Network (DNN) architectures Long Short-Term Memory (LSTM) and Gated Recurring Units (GRU) was presented by [2]. The authors used a 66-year S&P 500 time series dataset to develop a prediction model (1950 to 2016). To get the first-level prediction, the technique comprises passing the input data to the LSTM network and then transferring the output of the LSTM layer to the GRU layer. The proposed network outperformed earlier neural network techniques, achieving a Mean Squared Error (MSE) of 0.00098. Using information about the stock market, the model developed in [3] attempts to anticipate future prices. The accuracy of the model is evaluated by comparing it to comparable models and combining Adaline Neural Network (ANN) with modified Particle Swarm Optimization (PSO).

In [4], a hybrid intelligent model employing an Adaptive Network-based Fuzzy Inference System (ANFIS) and quantum-behaved particle swarm optimization was presented. The model is used to estimate market prices in the future. Due to the random character of stock prices, the hybrid BiLSTM-GRU model and traditional statistical models presented in [5] do not perform well. Utilizing Bi-LSTM and GRU networks, the proposed model outperforms other models. In [6], a hybrid stock prediction model was presented. The model comprises a noise-filtering approach, unique features, and a prediction based on machine learning. The technique of noise filtering is utilized to smooth historical stock price data by eliminating the cyclic component of the time series. The new characteristics are utilized to forecast the stock price. Traditional and deep machine learning techniques for prediction are investigated using machine learning-based prediction. The study of [7], discovered that utilizing deep learning techniques and machine learning models, such as the LSTM models, with hybrid multilingual sentiment data (data that has been translated from non-native English-speaking countries into English), is a more accurate method of stock market forecasting than utilizing other models with various data types. Research [8] examined the use of Principal Component Analysis (PCA), Empirical Mode Decomposition (EMD), and LSTM to forecast Thailand's stock values. Additionally, they utilized news sentiment analysis to see whether it would enhance the performance of the LSTM model. The study indicated that the suggested framework performed better than the baseline approaches and that news sentiment analysis enhanced the LSTM model's performance. The model described in [9] consists of an LSTM-GRU network and 25 features. Indicators of performance demonstrate that the suggested model is more accurate than competing models. Reference [10] discusses the difficulties of forecasting stock performance and the superiority of the suggested ensemble ARIMA-LSTM technique. In [11], a hybrid approach for predicting financial time series was presented. LSTM, Polynomial Regression (PR), and Chaos Theory make up the model. The existence of chaos in the model is evaluated and then modelled. For the first forecasts, the model time series will be put into LSTM. The sequence of mistakes obtained from LSTM predictions is PR-appropriate for error forecasting. The final hybrid model forecasts are produced by using error forecasts and original model projections. In his research, [12] contributes in two ways. First, he introduced a unique and resilient deep convolutional GAN architecture that serves in both generative and discriminative networks for stock price forecasting [19-28]. He also advised modifying the generator's loss function by adding additional phrases to improve prediction. Sentiment analysis is a technique for analysing research literature for the presence of underlying attitudes and emotions. It analyses sentences to determine if they convey optimistic or pessimistic sentiments, which may be used for the task of predicting the stock market. Positive news can be given a value of 1 and bad news a value of 0 to help gauge attitude. Stock market forecasts have also benefited from the application of deep learning models, which use existing data to spot patterns, and lexicon-based techniques, which examine the prevalence of particular words to ascertain their emotions.

The distance between market fluctuations and news feelings has also been closed with the use of text mining and natural language processing. Improved stock price forecasting is a primary motivation for the further development and investigation of these techniques [13-18]. Table 1 shows a Summary of the Sentiment analysis of previous research.

Through the above from previous studies, the hybrid model can provide better results compared to the unilateral model. By combining quantitative and qualitative data in one model and employing artificial intelligence techniques, a model with a better effect can be provided. The goal of demonstrating that GANs can also give better results for financial time series forecasting (in contrast to more conventional approaches) remains unresolved, despite the recent demonstration of promising results in addressing numerous complex situations (e.g., realistic image and video generation, image to-image, and text-to-image translation). This study uses GANs techniques by presenting a hybrid model for predicting stock markets. To the best of our knowledge, no previous studies have combined news and tweet sentiment analysis, historical stock price data, and cost functions for managing the flow of data into a unified stock prediction system. Sequence prediction issues, which include time series problems like stock market forecasting, need careful attention to temporal dependencies. To fill these knowledge gaps, we have attempted to 1) An approach to data processing is provided, focusing on the integration of quantitative and qualitative information into the suggested model. 2) Using GANs networks in processing time series data which is not been utilized previously in this context. 3) To create a high-performing model, we fused two powerful networks (RNN-GRU and DL-CNN). 4) Data and model output prediction flow management using several cost functions (Sigmoid, ReLU, and Cross Entropy).

Table 1: Summary of Sentiment analysis pervious research

Study	Data	Objective	Techniques	Result
[13]	News sentiments and historical stock prices	Predict the stock market	Random forest method	Sentiment classified
[14]	Financial news	Labelling of financial news	ML algorithms and lexicon-based	Noun-verb approach yields the best results
[15]	Financial news	Predict stock trend prediction	LSTM	Select useful indices of Stock
[16]	News sentiments	Predict the stock price	NLP	Improve the performance of the stock market
[17] [18]	Stock news headlines	Stock movement prediction	Vander, TextBlob, and RNN	BERT and RNN were much more accurate

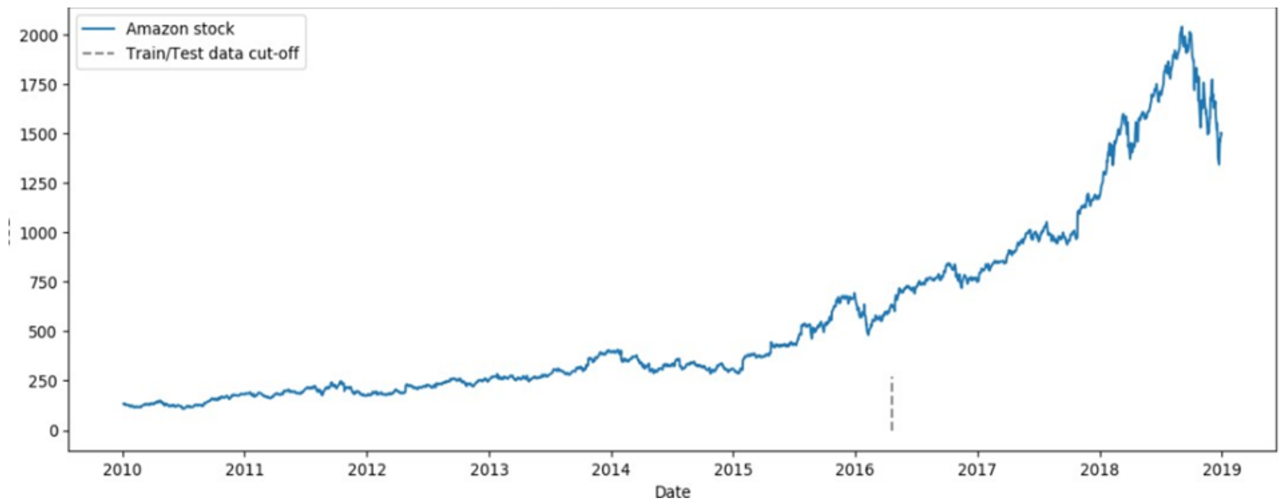

Figure 1: Amazon Stock Price

3 PROBLEM STATEMENT:

- How many different techniques is used to able to predict of stock?
- Which techniques are the best?
- Is it better to try to predict Stock prices or Stock price movement and use trading strategies?
- Can we tell when we can buy or sell the stock? What metrices can we use for different models?

4 WORK DONE & OBSERVATIONS

I predicted different models of the time series OHLCV data, performed feature extraction, hyper parameter tuning and trained on ARIMA, Fourier, LSTM, LSTM with sentimental analysis and GAN models. Then I focused on stock price movement instead of stock prices because I found that it's more accurate to predict them.

I took historical Amazon data from Yahoo. API and then performed feature generation on it, ARIMA model, Fourier model. Then performed LSTM on said model. Put up the data in time series form and split between train and test seen below.

Amazon data peaks around 2015 and after. Most of this data is in the testing set and training set does not have peak value data. I have predicted this could be a problem while dealing with conventional LSTM.

4.1 Feature Generation:

Following technical indicators were generated other than OHLCV data:

Bollinger bands: Bollinger Bands is used to define the prevailing high and low prices in a market to characterize the trading band of a financial instrument or commodity. Bollinger Bands are a volatility indicator. Bands are consisting of **Moving Average (MA)** line, a upper band and lower band. The upper and lower bands are simply MA adding and subtracting standard deviation.

4.1.1 EMA: Exponential Moving Average. **It** is a better version of a simple moving average that doesn't have SMAs lag. Moving averages just average out the data for a given time so we know how the company's closing price are trending for a given number of days. example for 4 days is price was 22,23 ,45,1 (the company crashed on 4th day) the average would be 23. Now 23 is a below average value so it gives us an idea that 45 was indeed just a fluke and that in fact the company was always making losses, EMA is calculated as:

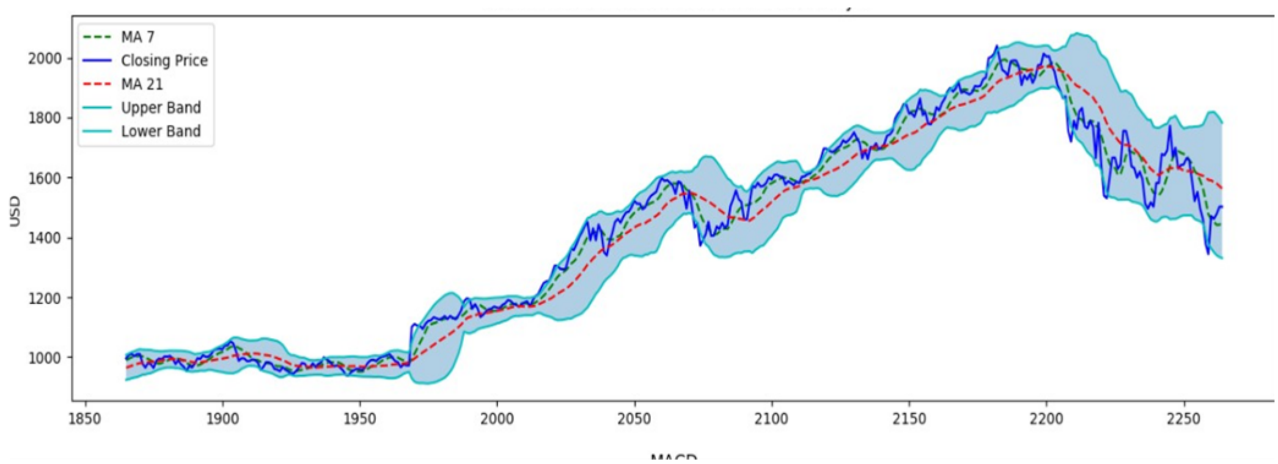

Figure 2: Technical Indicators for Amazon – Last 400 days

EMA(t) EMA(t0) =(1−α) EMA(t−1) +α p(t)=p(t0)
Where: α **=1L+1 and length of window is** α **=2M**

I used the ewm (exponential weighted mean) function to calculate ema.

4.1.2 Momentum: Momentum is perhaps the simplest and easiest oscillator (financial analysis tool) to understand and use. It is the measurement of the speed or velocity of price changes, or the rate of change in price movement for a particular asset.

The formula for momentum is: **Momentum=V−Vx**
Where**: V=Latest price Vx=Closing price x=Number of days ago**

while generating them other features that got generated were: 20SD (Standard Deviation of 20 days), Upper Band, Lower band, moving average of 7 days and 21 days and exponential moving average of 26 and 12 days.

Plot of technical indicators over days where they started peaking. They also peak around the days where Amazon saw a huge growth. Then I generated ARIMA and Fourier models and decided to see if they can be used as features.

4.2 ARIMA

4.3 Summary of ARIMA Model

- A good starting point for the AR parameter of the model may be 5 which we did.
- From the summary of the ARIMA we can see that most P-values are greater than 0.05 other than the last two. The model should be great!
- The difference between AIC and BIC is low so this indicates this is a good model
- Running the example, we can see that there is a positive correlation with the first 0-to-500 lags that is perhaps significant for the first 250 lags in the autocorrelation below:

5 FOURIER MODEL:

Use Fourier Transform in the spectral domain and reconvert it into time domain and plot with multiple components. The component which is closest to real values can be plotted.

In our case it is 100 components. Normalize the values i.e., do not keep spectral component values and generate the prediction data Fourier gives results very close to Closing price data as seen in plot (Figure 6):

6 SIMPLE MOVING AVERAGE:

I used the simple moving average by creating a lookback window and then ran it on the data. I was able to get a good model just as expected from SMA (Figure 7).

7 EXPONENTIAL MOVING AVERAGE:

EMA is a great model for this dataset. Ideally the pattern of the True data should have been followed in the prediction model. I coded from range to 1 to N-1 and put all the averaged values in the running mean. I used dense as 0.5 and then multiply it with the running average (Figure 8).

It predicts the predicted values after performing EMA on the dataset formula of which has been given above.

8 FEATURE IMPORTANCE USING XGBOOST:

Using XGBoost I found which features would make be the best for prediction. These are plotted below.

As seen above Figure 9, Open Adj Close, EMA and high and Low are great indicators. Others are ma7, ma21 and 12ema.

LSTM model to predict stock prices using 1 feature. I ran Open training data for 100 epochs and tried to predict Open with it. This is more like a regression problem so the metrics I used were mse and mean absolute error, not accuracy. The results weren't all that great since there was some overfitting and I normalized the data and did not perform hyperparameter tuning.

```
                              ARIMA Model Results
==============================================================================
Dep. Variable:                D.Close   No. Observations:                 2264
Model:                 ARIMA(5, 1, 0)   Log Likelihood               -9244.973
Method:                       css-mle   S.D. of innovations             14.361
Date:                Mon, 12 Aug 2019   AIC                          18503.947
Time:                        16:38:35   BIC                          18544.021
Sample:                             1   HQIC                         18518.569

=================================================================================
                    coef    std err          z      P>|z|      [0.025      0.975]
---------------------------------------------------------------------------------
const             0.6074      0.291      2.086      0.037       0.037       1.178
ar.L1.D.Close    -0.0270      0.021     -1.287      0.198      -0.068       0.014
ar.L2.D.Close    -0.0002      0.021     -0.008      0.993      -0.041       0.041
ar.L3.D.Close    -0.0293      0.021     -1.399      0.162      -0.070       0.012
ar.L4.D.Close    -0.0431      0.021     -2.055      0.040      -0.084      -0.002
ar.L5.D.Close     0.0631      0.021      2.956      0.003       0.021       0.105
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1           -1.3201           -0.9742j            1.6406           -0.3988
AR.2           -1.3201           +0.9742j            1.6406            0.3988
AR.3            0.6678           -1.5865j            1.7213           -0.1866
AR.4            0.6678           +1.5865j            1.7213            0.1866
AR.5            1.9877           -0.0000j            1.9877           -0.0000
-----------------------------------------------------------------------------
```

Figure 3: ARIMA Model Results

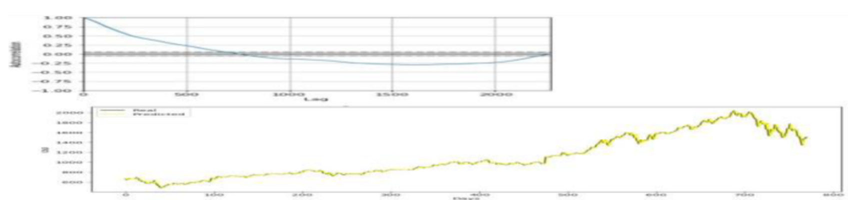

Figure 4: ARIMA Model on Amazon Stock

8.1 MAE was: 0.167

This means the average difference between input and output for all 2300 datapoints is 0.167. However, the value is for the days here so the MAE here is bad. (2350 length of dataset will be denominator. Difference between actual and predicted values should be so small that such a large denominator dividing the difference should put MAE in rage of 10^-3 ie 0.00then digits. Since MAE is 167.something*10^-3(0.167) difference is high.The prediction plot looked like this:

8.2 LSTM Model Using Multiple Features:

So, I tried more features (5 features) and I tried to predict closing prices with them. I encoded and normalized the 5 features and the prediction plot looked something like this: Again, this was a regression problem so the metrics I used where MSE and MAE and not accuracy.

MAE here was 0.016 during training and the plot continues the trend of the training data but I am not satisfied.

MAE calculation is sum(Y-X)/total data points = (2000-1980)/2350 here 2000 is the day with highest closing price. 1980 is the day with highest closing price in training set and 2350 will be the length of the data,

My mistakes with the 1st LSTM models were:

- Encoding and normalization lost a lot of precious data. We have only 2000 days of data and 1500 days of training data

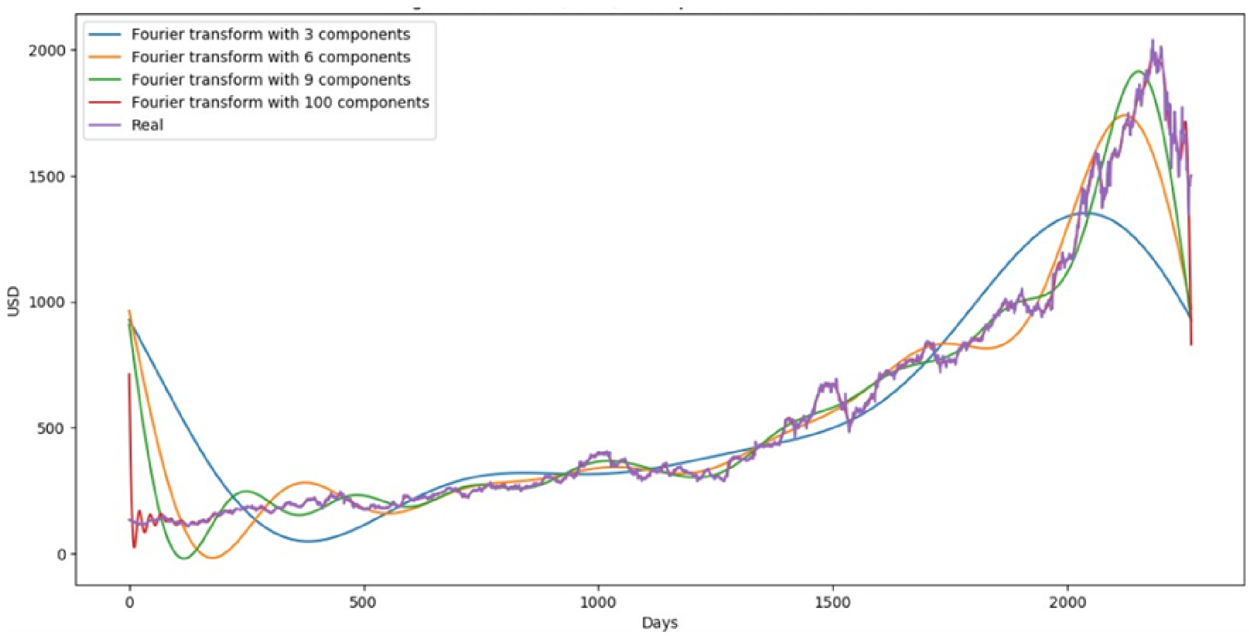

Figure 5: Amazon (close) stock price & Fourier Transforms

where every data point matters. So, no extensive encoding and normalization.

- I tried to predict stock prices as a regression problem but di not do any hyperparameter tuning. Adding linearity in a neural network will produce flat results.
- Instead of feeding direct stock prices I should feed in stock price movements by creating a window that keeps all similar stock price values in one window and feeds in each window as a datapoint.

8.3 LSTM PREDICTING STOCK PRICE MOVEMENT:

I created a next 'Vanila' LSTM model to predict only Closing stock prices and got a good prediction model.

- Ran a window over close data to convert closing prices into closing price movements.
- Normalized the data and split into train test and validation.
- Ran the model and saw to that it does not overfit.
- Got a MAE of 0.0076 lowest yet in any model
- Plotted a good prediction model:

Next LSTM models I will focus on hyperparameter tuning.

Since AMAZON dataset has only 2000 days 1500 days of which are kept for training and 500 for testing in which 500 testing days are the days where Amazon saw its exponential growth the models especially for LSTM will not show further improvement than this. So, I decided to move to a GE dataset that spans from 1970 to 2010s and gives more data to play around with.

- -I cleaned up the data.
- -Normalized, ran a window to get price movements and split into training and testing datasets -Performed hyperparameter tuning. Prediction plots:

The blue ticks indicate when the price movement changes. As you can see it is mostly making the right prediction for the price movement. If you are a trader this gives a good idea when to short your stock or invest more. The next plot gives a less confusing version:

9 LSTM SENTIMENTAL ANALYSIS MODELS:

There are tons of papers talking about LSTM and Twitter sentiments. I went with Reddit sentiments, news and Amazon data and drew in polarity between them and stock prices.

Using this I ran a LSTM model which used the polarity of the sentiment and prices (run a code to encode both) and then I predicted the test outputs. After hyperparameter tuning I got this model. The blue dots indicate the areas where price movement is present and where you can buy/sell the stock. (sell when lower buy when higher).

The prediction vs Actual plot. The predicted price follows the same structure as true but with less magnitude. All in all, the sentimental analysis gave a great model for the limited Amazon dataset and I could say when to buy or sell stock.

10 GAN MODEL:

The GAN model does not work as the same regression problem. I used the time series data as generator and a CNN as a discriminator. I made a confusion matrix that says how close I am to detecting if the price is going up or down.

Fourier	ARIMA	Close
133.899994	NaN	133.899994
134.690002	NaN	134.690002
132.250000	NaN	132.250000
130.000000	NaN	130.000000
133.520004	NaN	133.520004
130.309998	NaN	130.309998
127.349998	NaN	127.349998
129.110001	NaN	129.110001
127.349998	NaN	127.349998
127.139999	NaN	127.139999
127.610001	NaN	127.610001
125.779999	NaN	125.779999
126.620003	NaN	126.620003
121.430000	NaN	121.430000
120.309998	NaN	120.309998
119.480003	NaN	119.480003
122.750000	NaN	122.750000
126.029999	NaN	126.029999
125.410004	NaN	125.410004
118.870003	NaN	118.870003

Figure 6: Fourier & Closing Price

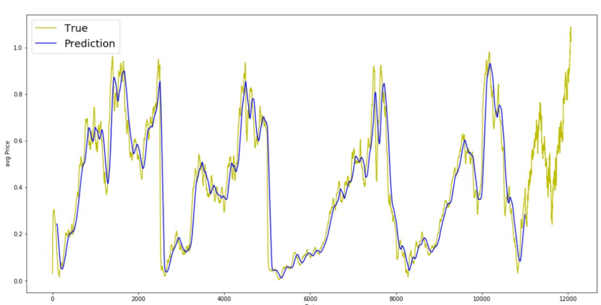

Figure 7: Prediction using SMA

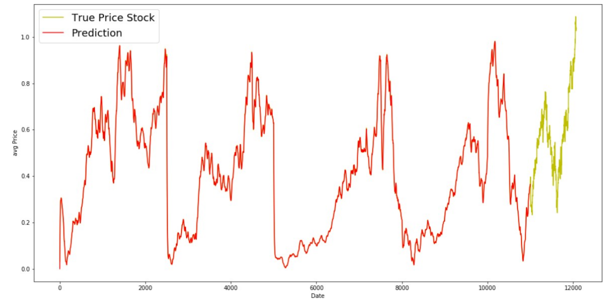

Figure 8: Predictions Using EMA

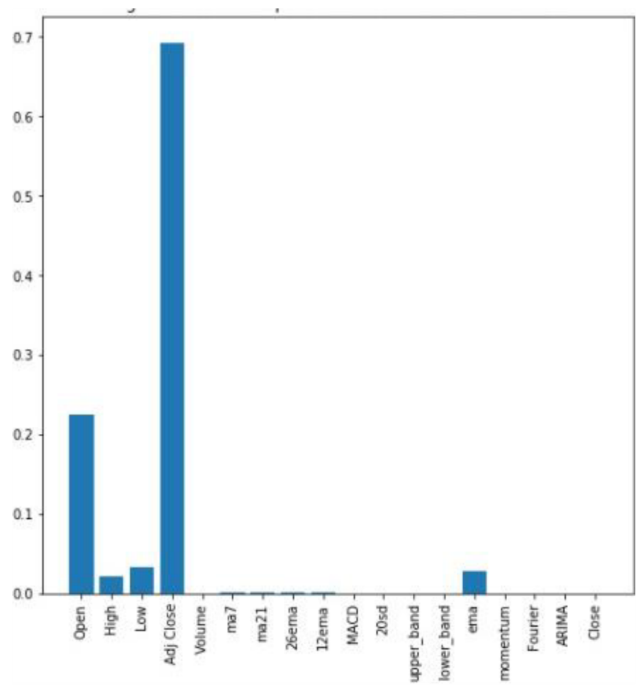

Figure 9: Feature Importance Using XGBoost

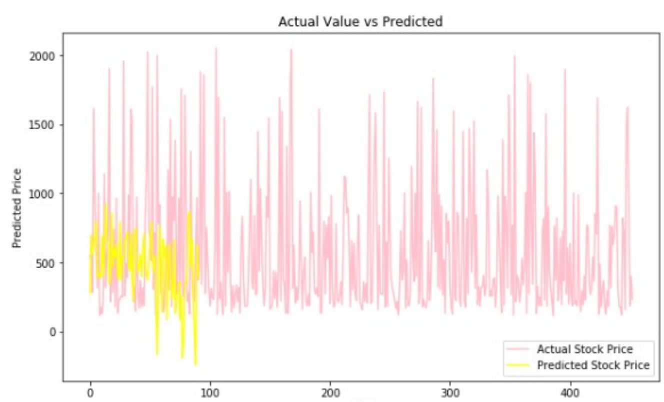

Figure 10: Prediction Using MAE

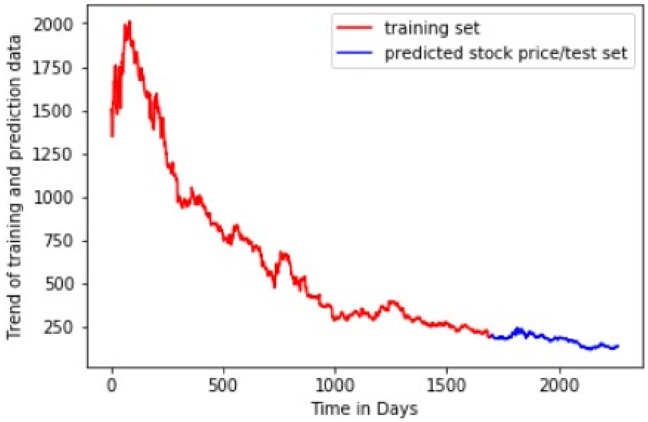

Figure 11: Prediction Using LSTM

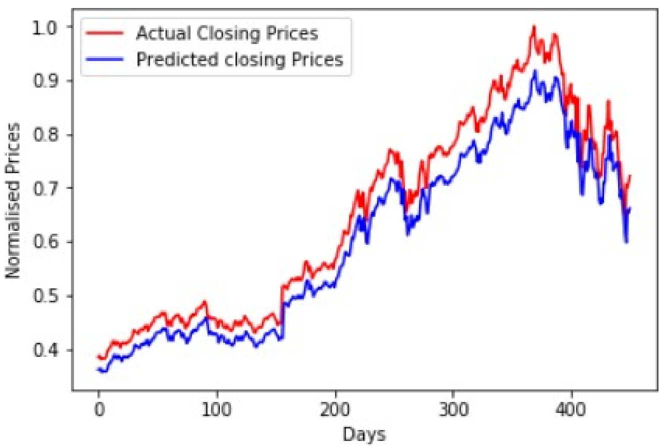

Figure 12: LSTM PREDICTING STOCK PRICE MOVEMENT

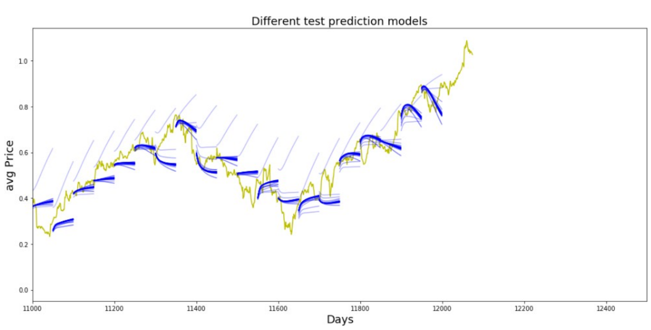

Figure 13: Different test prediction model

Chances od Down price being predicted as Down is 93% and Up and UP is 87% which is quite good.

11 LSTM, GRU USING ACCURACY AS A METRIC:

- I used MAE as a metric of measure. However, I then changed to using accuracy as a measure by changing the regression problem into classification where I found how I could predict datapoints of the closing price by creating EMA for the given dataset.
- I then used this dataset to split into test, train and validate sets.
- I stacked the OHLCV data and used that as X and closing prices as Y hence the classification problem.
- Then I predicted the accuracy.

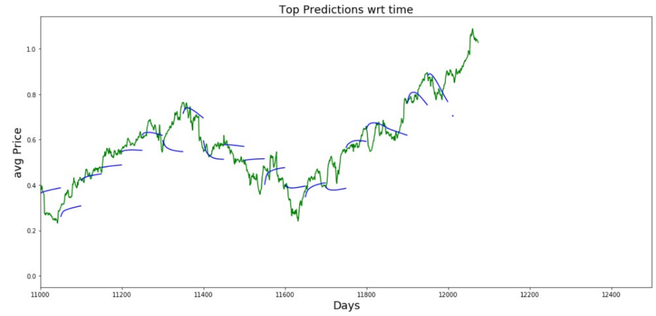

Figure 14: Top Predictions wrt time

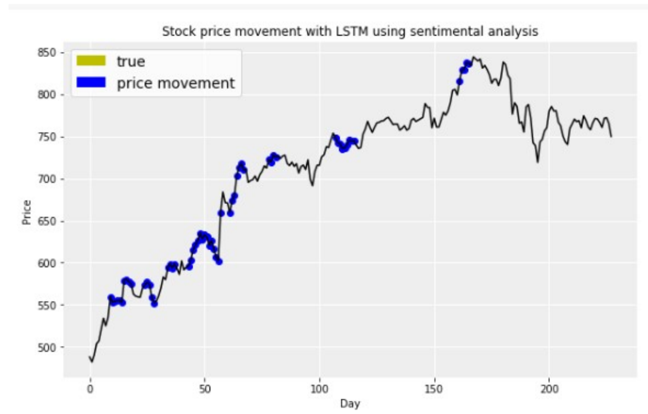

Figure 15: Stock Price movement with LSTM using sentimental analysis

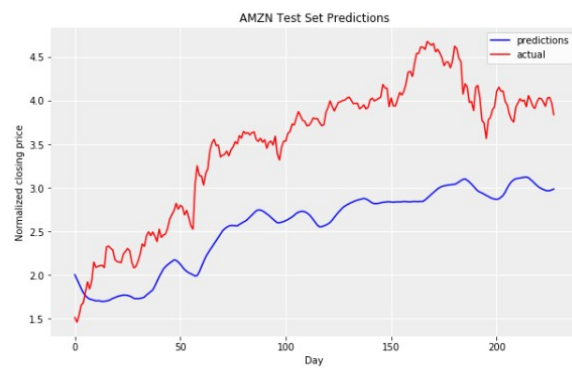

Figure 16: AMZN test set predictions

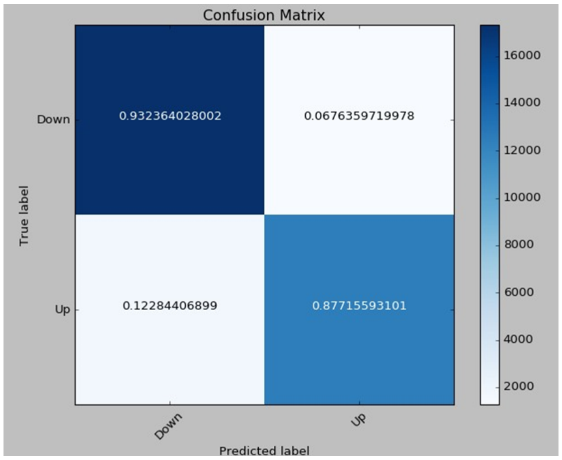

Figure 17: Confusion Matrix

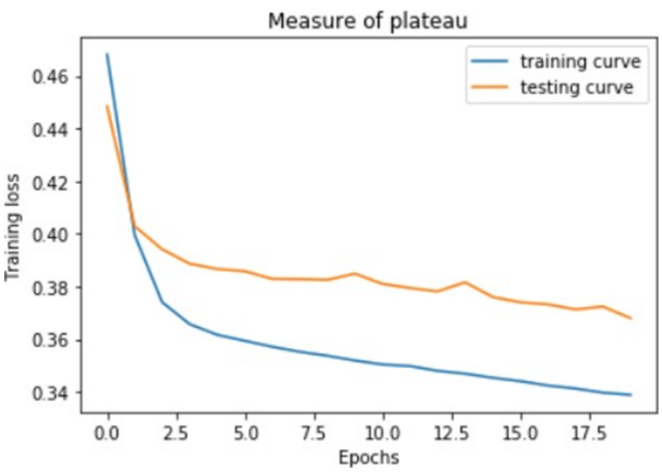

Figure 18: Training and test set accuracy curve

- I used hyperparameter tuning for the LSTM model and ran a GRU cell with some hyperparameter too.

Results of Best LSTM model

This model gave an accuracy of 68.9% i.e., it was able to predict the right closing prices 68.9% of the time.

REFERENCES

[1] J. J. Wang, J. Z. Wang, Z. G. Zhang and S. P. Guo, "Stock index forecasting based on a hybrid model," *Omega*, vol. 40, no. 6, pp. 758–766, 2012

[2] M. A. Hossain, R. Karim, R. Thulasiram, N. D. Bruce and Y. Wang, "Hybrid deep learning model for stock price prediction,"in *2018 IEEE Symp. Series on Computational Intelligence (SSCI)*, Bangalore, India, IEEE, pp. 1837–1844, 2018.

[3] M. Sedighi, H. Jahangirnia, M. Gharakhani and S. Farahani Fard, "A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine," *Data*, vol. 4, no. 2, pp. 75, 2019.

[4] A. Bagheri, H. M. Peyhani and M. Akbari, "Financial forecasting using ANFIS networks with quantum behaved particle swarm optimization," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6235–6250, 2014.

[5] M. Karim, M. Foysal and S. Das, "Stock price prediction using Bi-LSTM and GRU-based hybrid deep learning approach," *in Proc. of Third Doctoral Symp. on Computational Intelligence*, Singapore, Springer, pp. 701–711, 2023

[6] Q. M. Ilyas, K. Iqbal, S. Ijaz, A. Mehmood and S. Bhatia, "A hybrid model to predict stock closing price using novel features and a fully modified hodrick–Prescott filter," *Electronics*, vol. 11, no. 21, pp. 3588, 2022.

[7] Y. L. Lin, C. J. Lai and P. F. Pai, "Using deep learning techniques in forecasting stock markets by hybrid data with multilingual sentiment analysis," *Electronics*, vol. 11, no. 21, pp. 3513, 2022.

[8] K. Srijiranon, Y. Lertratanakham and T. Tanantong, "A hybrid framework using PCA, EMD and LSTM methods for stock market price prediction with sentiment analysis," *Applied Sciences*, vol. 12, no. 21, pp. 10823, 2022.

[9] G. R. Patra and M. N. Mohanty, "An LSTM-GRU based hybrid framework for secured stock price prediction," *Journal of Statistics and Management Systems*, vol. 25, no. 6, pp. 1491–1499, 2022.

[10] S. Verma, S. Prakash Sahu and T. Prasad Sahu, "Ensemble approach for stock market forecasting using ARIMA and LSTM model," in *Proc. of Third Int. Conf. on Intelligent Computing, Information and Control Systems*, Singapore, Springer, pp. 65–80, 2022.

[11] M. Durairaj and K. M. BH, "Statistical evaluation and prediction of financial time series using hybrid regression prediction models," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 9, no. 4, pp. 245–255, 2021.

[12] A. Staffini, "Stock price forecasting by a deep convolutional generative adversarial network," *Frontiers in Artificial Intelligence*, vol. 5, pp. 1–16, 2022.

[13] A. Bhardwaj, Y. Narayan and M. Dutta, "Sentiment analysis for Indian stock market prediction using Sensex and nifty," *Procedia Computer Science*, vol. 70, pp. 85–91, 2015.

[14] A. Yadav, C. K. Jha, A. Sharan and V. Vaish, "Sentiment analysis of financial news using unsupervised approach," *Procedia Computer Science*, vol. 167, pp. 589–598, 2020.

[15] X. Jiawei and T. Murata, "Stock market trend prediction with sentiment analysis based on LSTM neural network," in *Int. Multiconf. of Engineers and Computer Scientists*, Hong Kong, pp. 475–479, 2019.

[16] P. Mehta, S. Pandya and K. Kotecha, "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning," *PeerJ Computer Science*, vol. 7, pp. e476, 2021. CSSE, 2023, vol.47, no.1 35

[17] T. B. Shahi, A. Shrestha, A. Neupane and W. Guo, "Stock price forecasting with deep learning: A comparative study," *Mathematics,* vol. 8, no. 9, pp. 1441, 2020.

[18] K. S. Rekha and M. K. Sabu, "A cooperative deep learning model for stock market prediction using deep autoencoder and sentiment analysis," *PeerJ Computer Science*, vol. 8, pp. e1158, 2022.

[19] Singh, U. P., Saxena, V., Kumar, A., Bhari, P., & Saxena, D. (2022, December). Unraveling the Prediction of Fine Particulate Matter over Jaipur, India using Long Short-Term Memory Neural Network. In *Proceedings of the 4th International Conference on Information Management & Machine Intelligence* (pp. 1-5).

[20] Kumar, A., Bhari, P. L., Singh, U. P., & Saxena, V. (2022, December). Comparative Study of different Machine Learning Algorithms to Analyze Sentiments with a Case Study of Two Person's Microblogs on Twitter. In *Proceedings of the 4th International Conference on Information Management & Machine Intelligence* (pp. 1-6).

[21] Saxena, V., Saxena, D., & Singh, U. P. (2022, December). Security Enhancement using Image verification method to Secure Docker Containers. In *Proceedings of the 4th International Conference on Information Management & Machine Intelligence* (pp. 1-5).

[22] Chauhan, M., Malhotra, R., Pathak, M., & Singh, U. P. (2012). Different aspects of cloud security. *International Journal of Engineering Research and Applications*, *2*, 864-869.

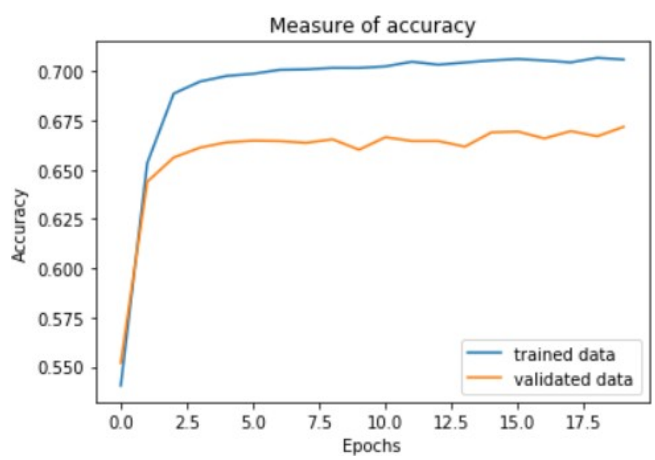

Figure 19: Training and Validation Set accuracy curve

[23] Mittal, A. K., Singh, U. P., Tiwari, A., Dwivedi, S., Joshi, M. K., & Tripathi, K. C. (2015). Short-term predictions by statistical methods in regions of varying dynamical error growth in a chaotic system. *Meteorology and Atmospheric Physics*, *127*, 457-465.

[24] Singh, U. P., Mittal, A. K., Dwivedi, S., & Tiwari, A. (2015). Predictability study of forced Lorenz model: an artificial neural network approach. *History*, *40*(181), 27-33.

[25] Singh, U. P., Mittal, A. K., Dwivedi, S., & Tiwari, A. (2020). Evaluating the predictability of central Indian rainfall on short and long timescales using theory of nonlinear dynamics. *Journal of water and Climate Change*, *11*(4), 1134-1149.

[26] Singh, U., Pathak, M., Malhotra, R., & Chauhan, M. (2012). Secure communication protocol for ATM using TLS handshake. *Journal of Engineering Research and Applications (IJERA)*, *2*(2), 838-948.

[27] Singh, U. P., & Mittal, A. K. (2021). Testing reliability of the spatial Hurst exponent method for detecting a change point. *Journal of Water and Climate Change*, *12*(8), 3661-3674.

[28] Tiwari, A., Mittal, A. K., Dwivedi, S., & Singh, U. P. (2015). Nonlinear time series analysis of rainfall over central Indian region using CMIP5 based climate model. *Climate Change*, *1*(4), 411-417.