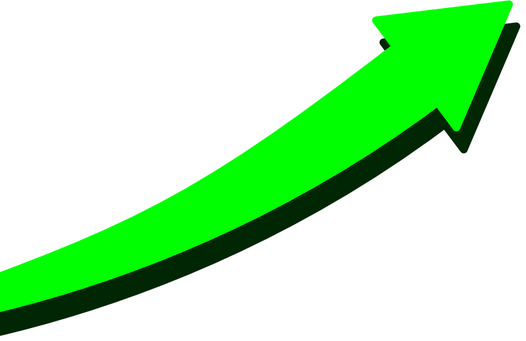
USE OF MACHINE LEARNING FOR Stock Market Prediction

Anjani Kumar

Shreyas M Honrao

Priyanshu Jogdand



INTRODUCTION

In this study, we endeavor to tackle the stock price prediction conundrum by harnessing the capabilities of three distinct machine learning models: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Linear Regression. These models leverage historical stock prices and pertinent features to formulate informed projections regarding future stock prices.

OBJECTIVE

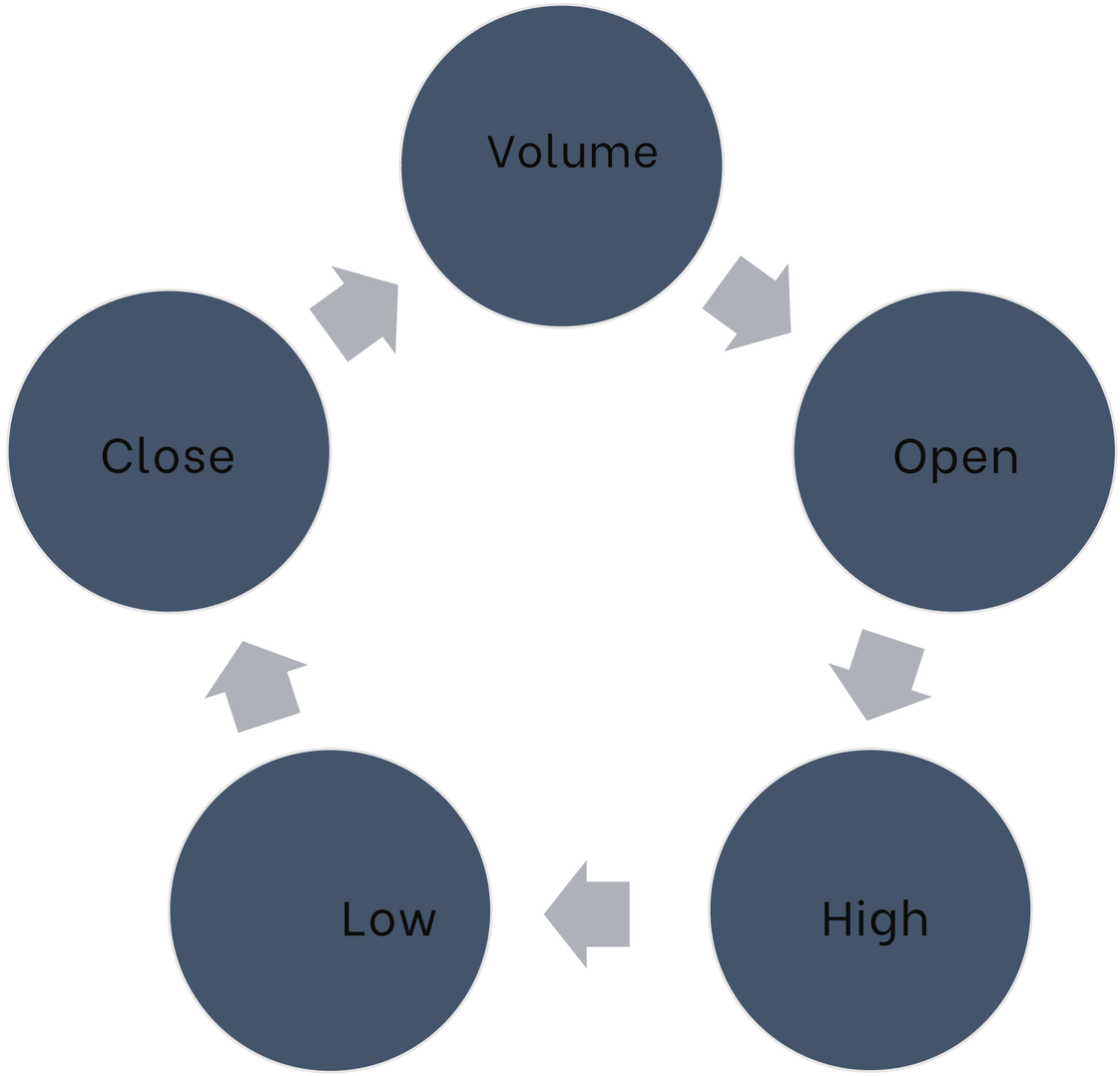
The primary objective of this research is to conduct a comprehensive comparative analysis of these models to determine their effectiveness in stock price prediction. Through meticulous evaluation of the performance of LSTM, GRU, and Linear Regression models, we aim to discern which model offers the highest predictive accuracy. By shedding light on the strengths and limitations of each model, we aim to contribute valuable insights to the domain of stock market forecasting.

Dataset

We have taken Dataset of Apple Inc which is traded in New York Stock Exchange(NYSE).

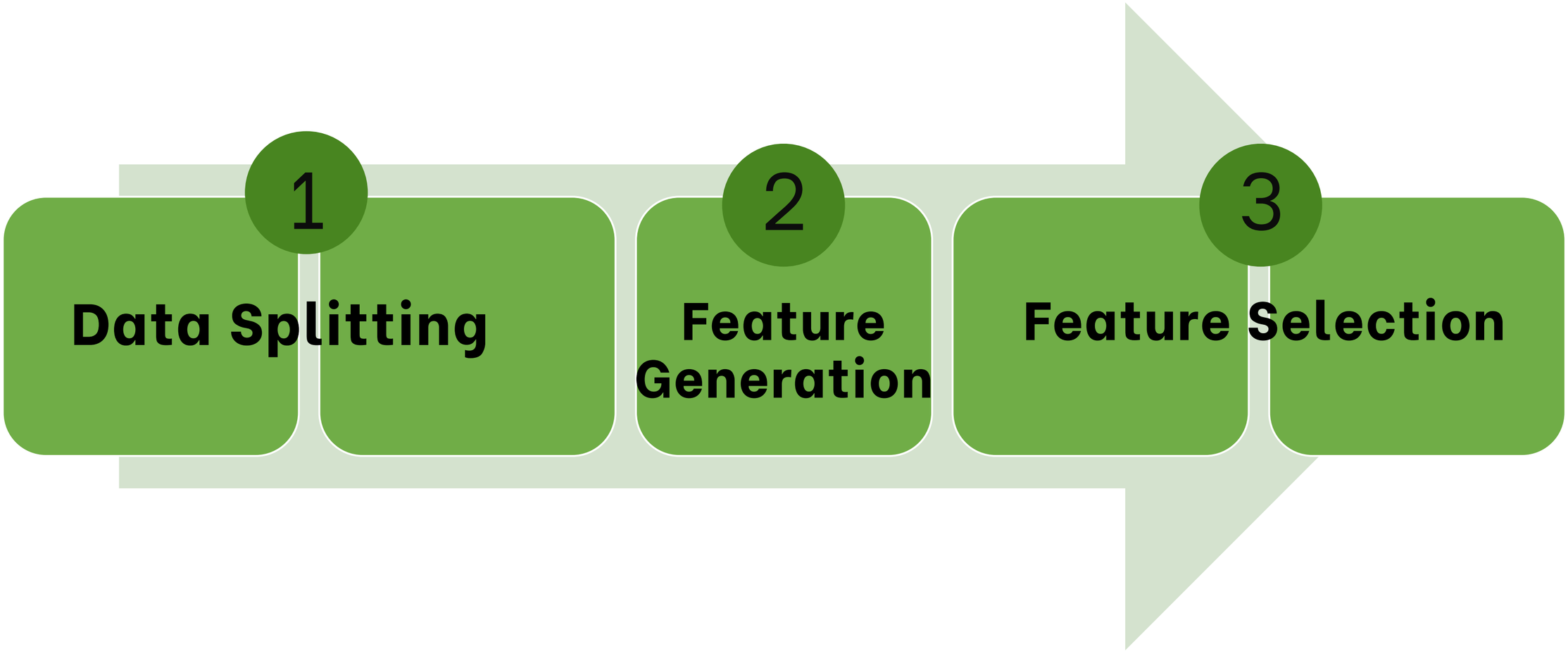
Dataset comprises data from 1 Jan 2006 to 1 Jan 2018

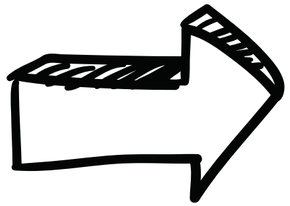
INPUT FEATURES IN THE DATASET



APPLE DATA SINCE 2006 TO 2018

DATA PIPELINE



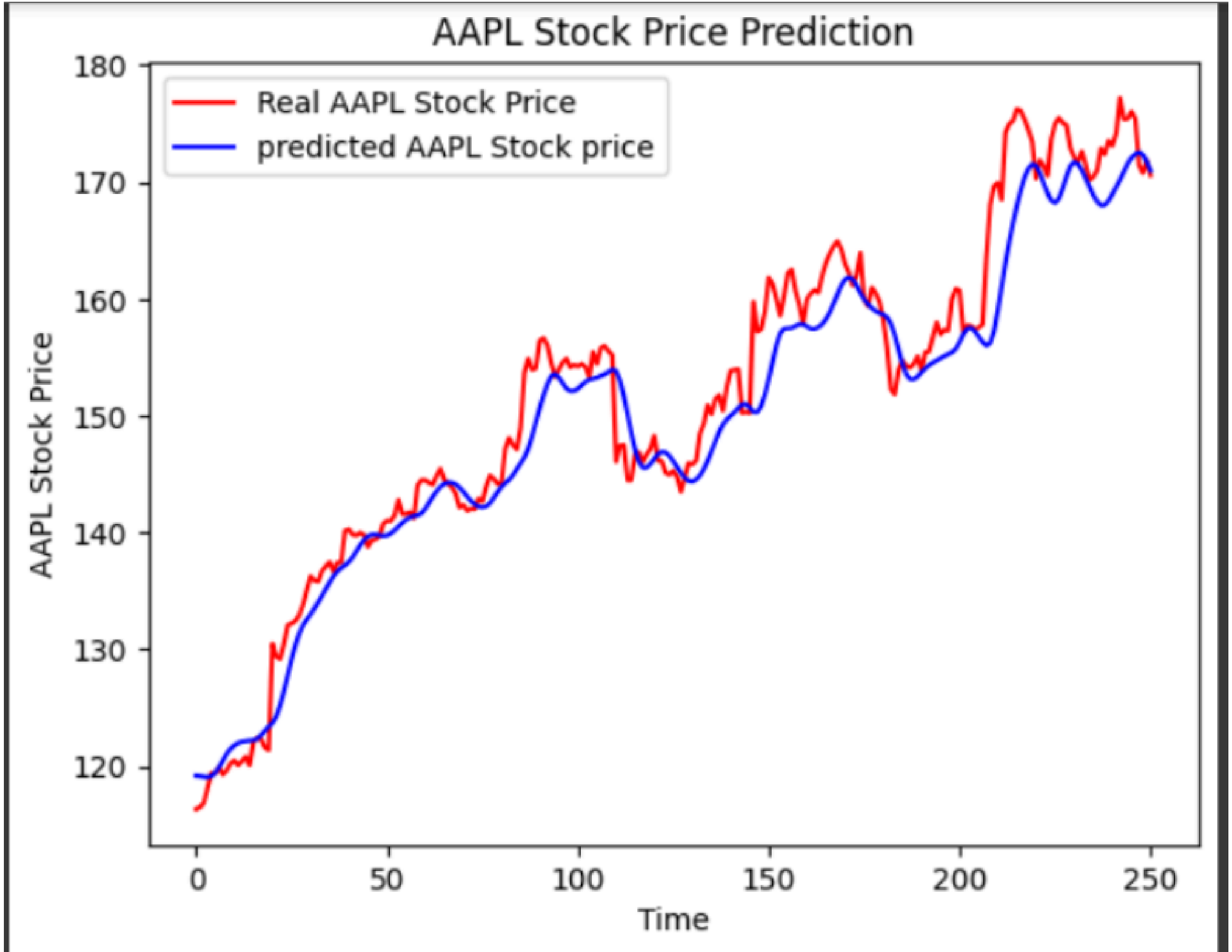


DATA PREPROCESSING AND SPLITTING

- We have used the method of Scaler Transform for Preprocessing.
- This method transforms each feature in the range of (0-1).
- Our data is splitted in 11:1 ratio.
- Training Data: 88.89%
- Testing Data: 11.11%
- After prediction inverse scaling is used to return the predicted prices into their original scale

LSTM MODEL

Overview	Building the model	Making predictions	Evaluating the model
RNN based Architecture used to vanish gradient descent problem and capture long term dependencies in the sequential data	LSTM model is build by Keras Sequential API. We have used 4 sequential layers and a dropout layer tfor regularization and finally a dense layer for the output Reason why we used 4 layers is for extracting complex features from the preprocessed data and also for avoiding overfitting	For prediction we are using previous 60 days of data to predict the next day's price. This is called Sliding Window approach which helps in increasing accuracy for the prediction Implemented loss function to check how well the model correlates with actual prices	We have evaluated our model based on 2 main aspects 1)Root mean Square error. 2) Stock returns on daily basis


AAPL Stock Price Prediction
Real AAPL Stock Price
predicted AAPL Stock price
AAPL Stock Price
Time
180
170
160
150
140
130
120
0
50
100
150
200
250

LSTM

GRU MODEL

Overview

GRU is also RNN based Architecture but it simplifies the LSTM architecture by combining the forget and input gates into a single update gate merging the cell and hidden states

Building the model

GRU model is build by Keras Sequential API.

We have used 4 sequential layers and a dropout layer for regularization and finally a dense layer for the output. In GRU we have used an activation function tanh to introduce non linearity for identifying complex patterns

Reason why we used 4 layers is for extracting complex features from the preprocessed data and also for avoiding overfitting

Making predictions

For prediction we are using previous 60 days of data to predict the next day's price. This is called Sliding Window approach which helps in increasing accuracy for the prediction

Also Used Adam Optimizer for adapting the learning rates of each parameter during training, allowing for faster convergence and improved performance on a wide range of problems.

Implemented loss function to check how well the model correlates with actual prices

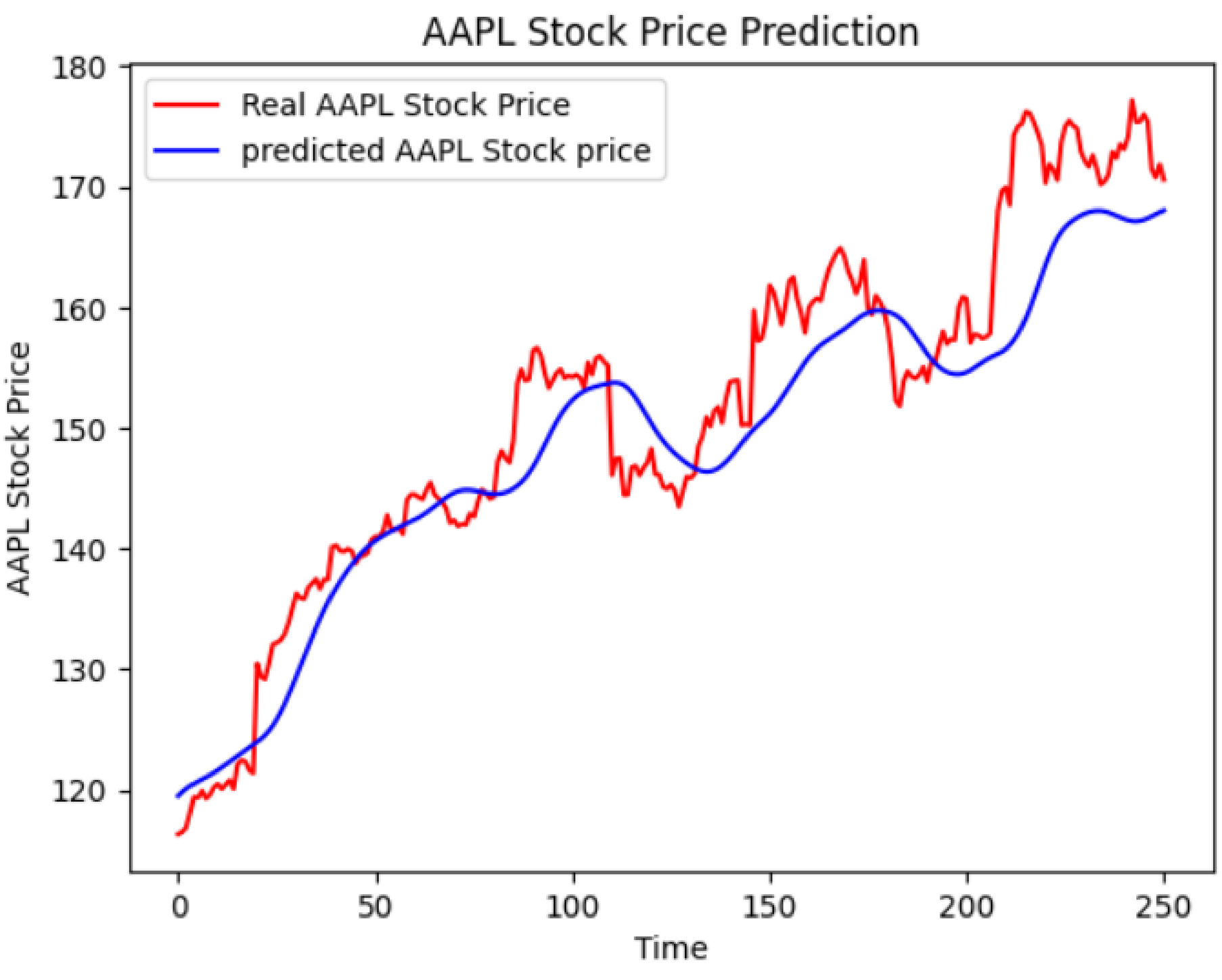
Evaluating the model

We have evaluated our model based on 2 main aspects

1)Root mean Square error.

.

2) Stock returns on daily basis


AAPL Stock Price Prediction
Real AAPL Stock Price
predicted AAPL Stock price
AAPL Stock Price
Time
180
170
160
150
140
130
120
0
50
100
150
200
250

GRU

LR MODEL

Overview

Linear Regression Model is stastical method that follows a relationship between a dependent variable(Future stcok price) and one or more independent variables(historical stock prices, volume).

Building the model

Start by defining dependent and independent variables

After training data is ready fit the model to minimize difference between actual and predicted values

Making predictions

Prediction are done on how model fits the linear curve for the whole dataset.

So for that Calculated performance of R^2(the proportion of variance of historical prices

If R^2 is high the model is fitted perfectly

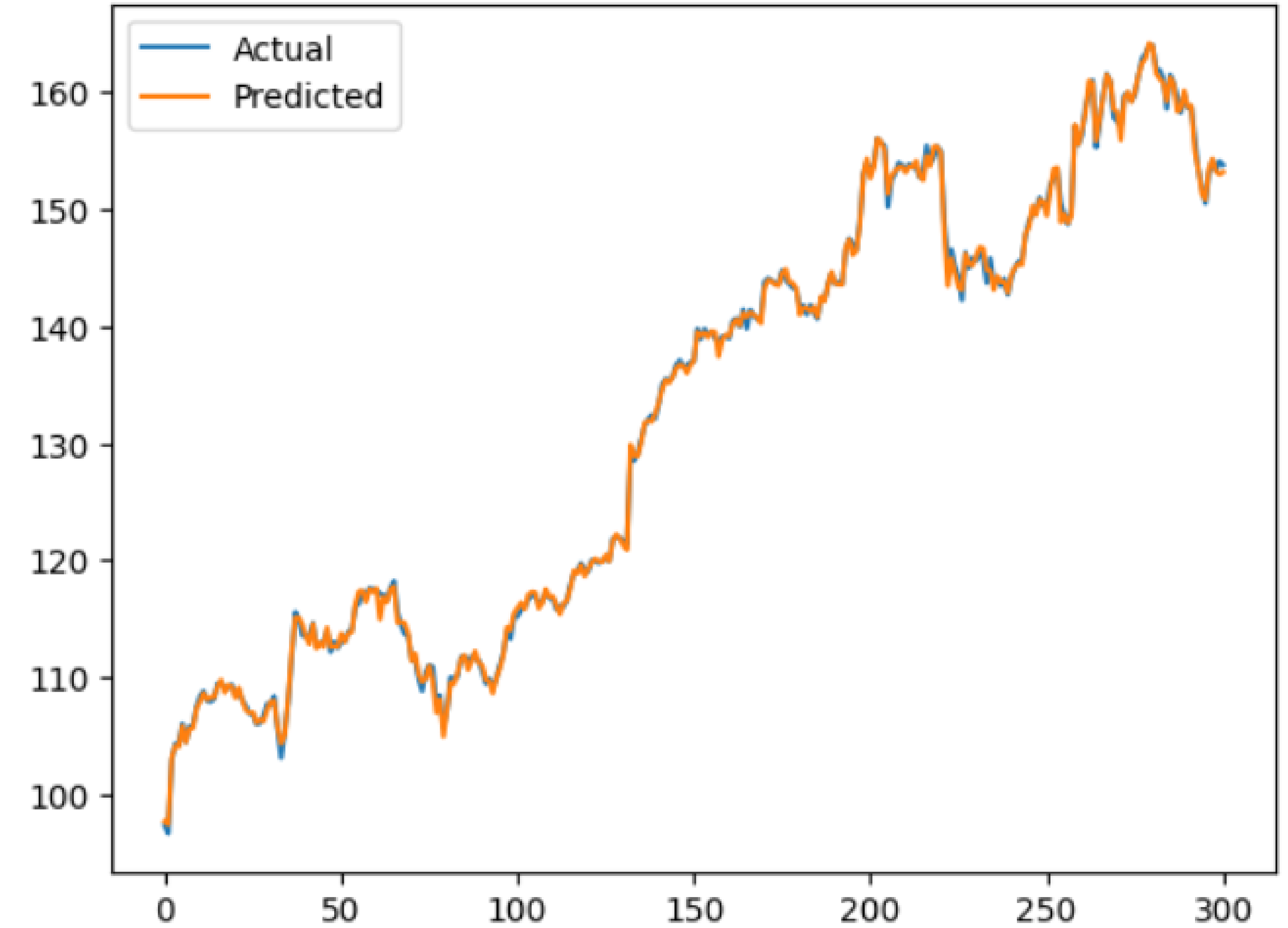
Evaluating the model

We have evaluated our model based on 3 main aspects

1)Root mean Square error.

2) Stock returns on daily basis

3) Mean Absolute Error


Actual
Predicted
160
150
140
130
120
110
100
0
50
100
150
200
250
300

LR

PERFORMANCE METRICS

Root Mean Square Error	RMSE is crucial in stock market prediction, measuring the average deviation between predicted and actual values. Lower RMSE suggests better alignment with real market behavior, aiding investors in selecting more accurate forecasting models for informed decisions.
Stock Returns on daily basis	Stock daily returns reflect the percentage shift in a stock's price between consecutive trading days, crucial for financial analysis tasks like risk assessment, portfolio optimization, and performance evaluation.

CONCLUSION

We did a comaprision between LSTM ,GRU and Linear Regression model . Our dataset was huge and we analyzed the results of all three models.

.

LSTM and GRU dont have computational power to analyze large datasets whereas LR model can fit better on huge datasets

Conclusion LR worked better in our case though not necessary always. (Our R^2 performance, RMSE, Stock Returns (actual and predicted accuracy) shows that).

Github Repository

Link for GITHUB Repository:

https://github.com/ShreyasHonrao/Computational-Intelligence-Project

REFERENCES

- https://link.springer.com/article/10.1007/s13042-019-01041-1
- https://www.hindawi.com/journals/sp/2021/4055281/
- https://search.proquest.com/openview/6d3e7bbf0c63bea87b9e2051f94b3e71/1?pq-origsite=gscholar&cbl=236248
- https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2
- https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

THANK YOU