Stock Market Prediction

USING

MACHINE LEARNING

Mohit Goyal
Bhavishya Aggarwal
Param Ahuja
Anand Mishra
Pratham Nahta

Agenda

- Abstract
- Introduction
- Existing Methods
- Proposed System
- System Design
- Implementation
- Results & Discussions
- Performance
- Future Enhancement
- Conclusion

Abstract

- Time series forecasting has been widely used to determine the future prices of stock, and the analysis and modelling of finance time series importantly guide investors' decisions and trades

- This work proposes an intelligent time series prediction system that uses various algorithms, for the purpose of predicting the stock prices

- The system has a graphical user interface and functions as a stand-alone application.

- The proposed model is a promising predictive technique for highly non-linear time series, whose patterns are difficult to capture by traditional models.

Introduction

- Financial markets are highly volatile and generate huge amounts of data daily
- It is the most popular financial market instrument and its value changes quickly
- Stock prices are predicted to determine the future value of companies' stock or other financial instruments that are marketed on financial exchanges
- However, the stock market is influenced by many factors such as political events, economic conditions and traders' expectation

What is Machine Learning

Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed

More formally, it can defined as,

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E

Example: playing checkers.

*E = the **experience** of playing many games of checkers*

*T = the **task** of playing checkers.*

*P = the **probability** that the program will win the next game.*

Machine Learning in Stock Prediction

- The field of Machine Learning is vast and plays a key role in a wide range of critical applications.
- The concept of **Support Vector Machines (SVM)** have advanced features that are reflected in their good generalization capacity and fast computation.
- Predicting the stock market involves predicting the closing prices of a company's stock for any given number of days ahead.
- SVMs can be used to perform Linear Regression on previous stock data to predict the closing prices using Time series forecasting and other optimization algorithms

Existing Methods

Time series forecasting consists of a research area designed to solve various problems, mainly in the financial area:

- LSTM
- NAIVE
- ARIMA
- LINEAR MODEL
- DENSE MODEL

Disadvantages

- The existing system focuses on the stock price market of a specific stock exchange, but does not generalize for other markets worldwide.
- The system does not allow the import of raw data directly
- The existing system cannot be used to analyze multi-variate time series
- Lastly, the system does not have a user-interface which can be distributed as a web app to users for personal use

Proposed System

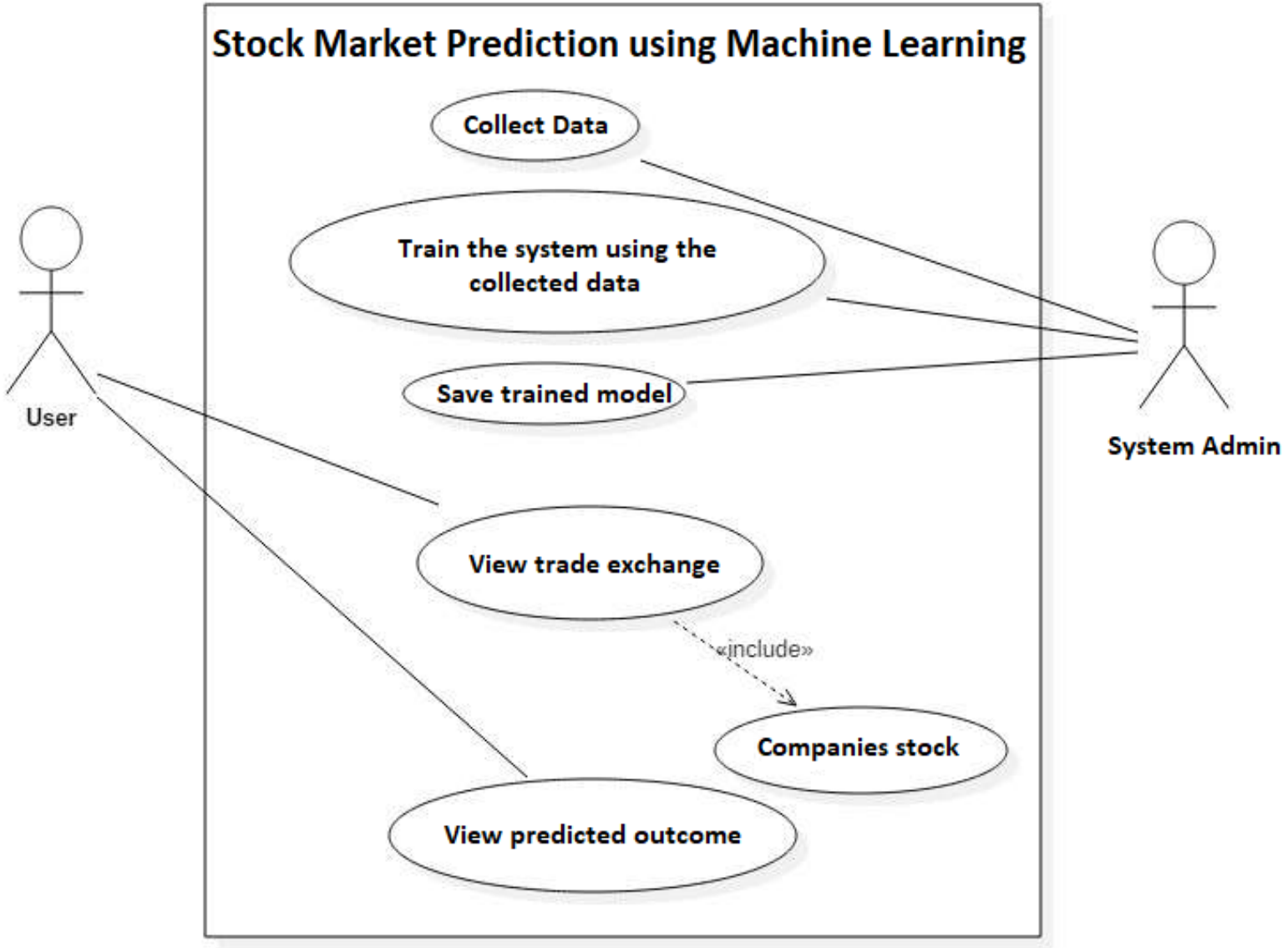
- To generalize the application of the existing system, our work uses the system to estimate other stocks in similar emerging markets and mature markets
- The system can be extended to analyze multivariate time series data and import raw dataset directly
- Profit can be maximized even when the corporate stock market is has lower value
- The development of a web-based application has been considered to improve the user-friendliness and usability of the expert system.

Architecture

System Design

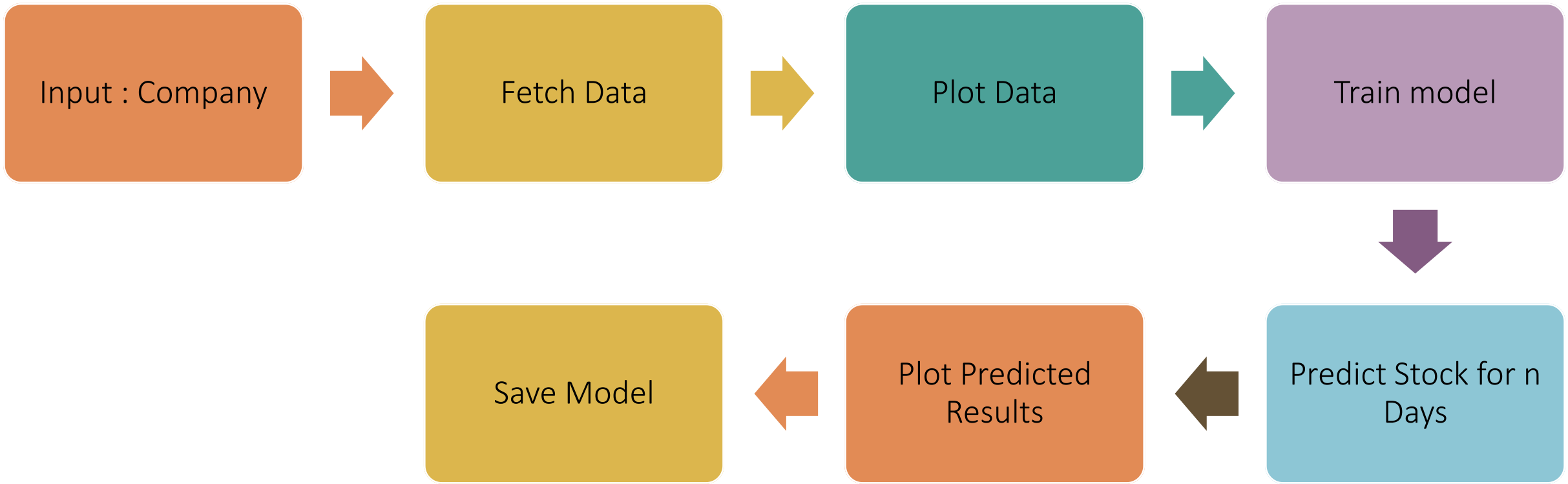
USE CASE DIAGRAM

1. Data is initially collected from online sources or the stock exchange
2. The data is then used to train the system
3. Trained model is saved
4. User views the trade exchange and stock of a company
5. Using the model, closing prices are predicted



Architecture

Data Flow Diagram

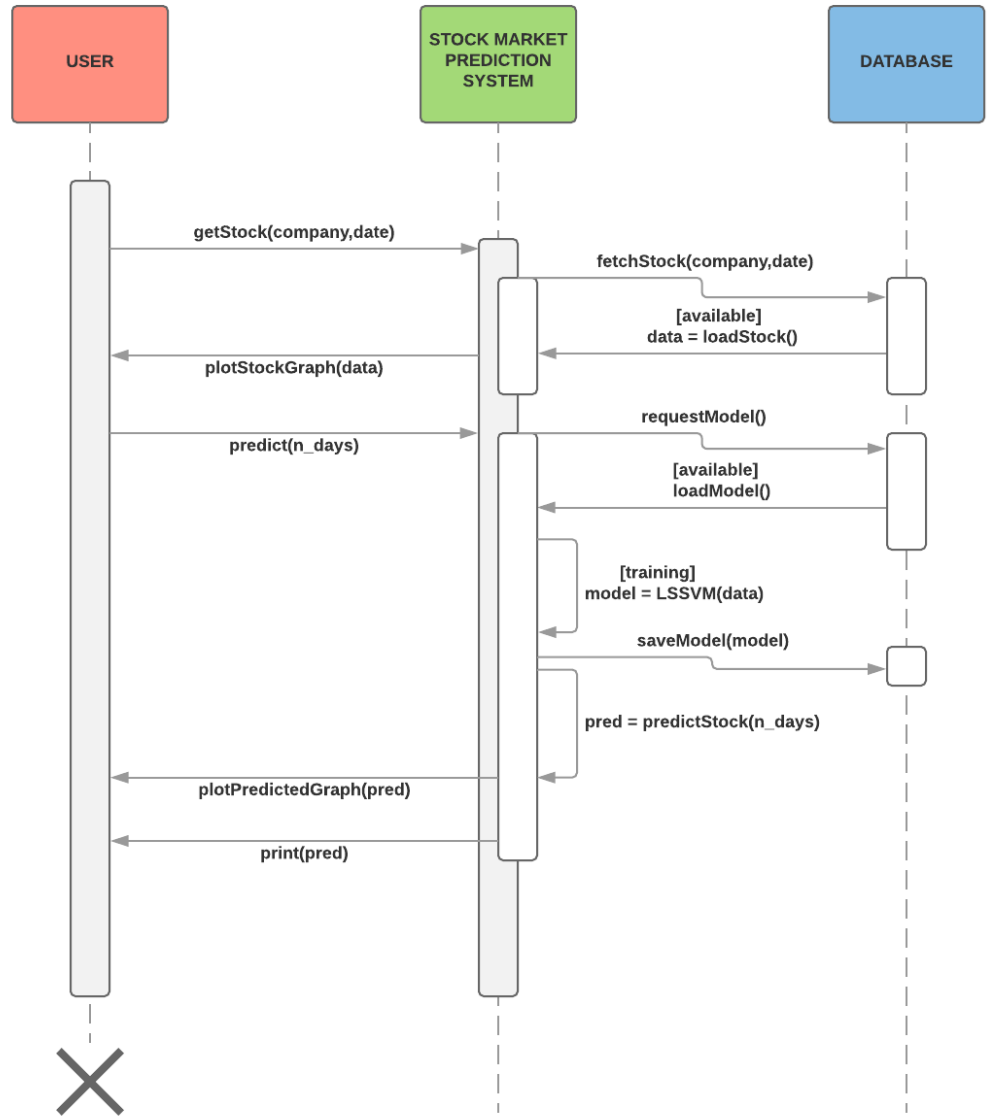


Architecture

Sequence Diagram

SEQUENCE DIAGRAM

1. User visits the website/webapp
2. Previously saved model is loaded
3. User requests for a company's stock data
4. He requests for prediction to be made
5. The Stock Market Prediction System trains a model using the data from the database
6. The model is saved for further use and closing price is predicted
7. Result is displayed along with graph



METHODOLOGY

P Python Language

Python is a rich language for Data Science and AI

A Algorithm

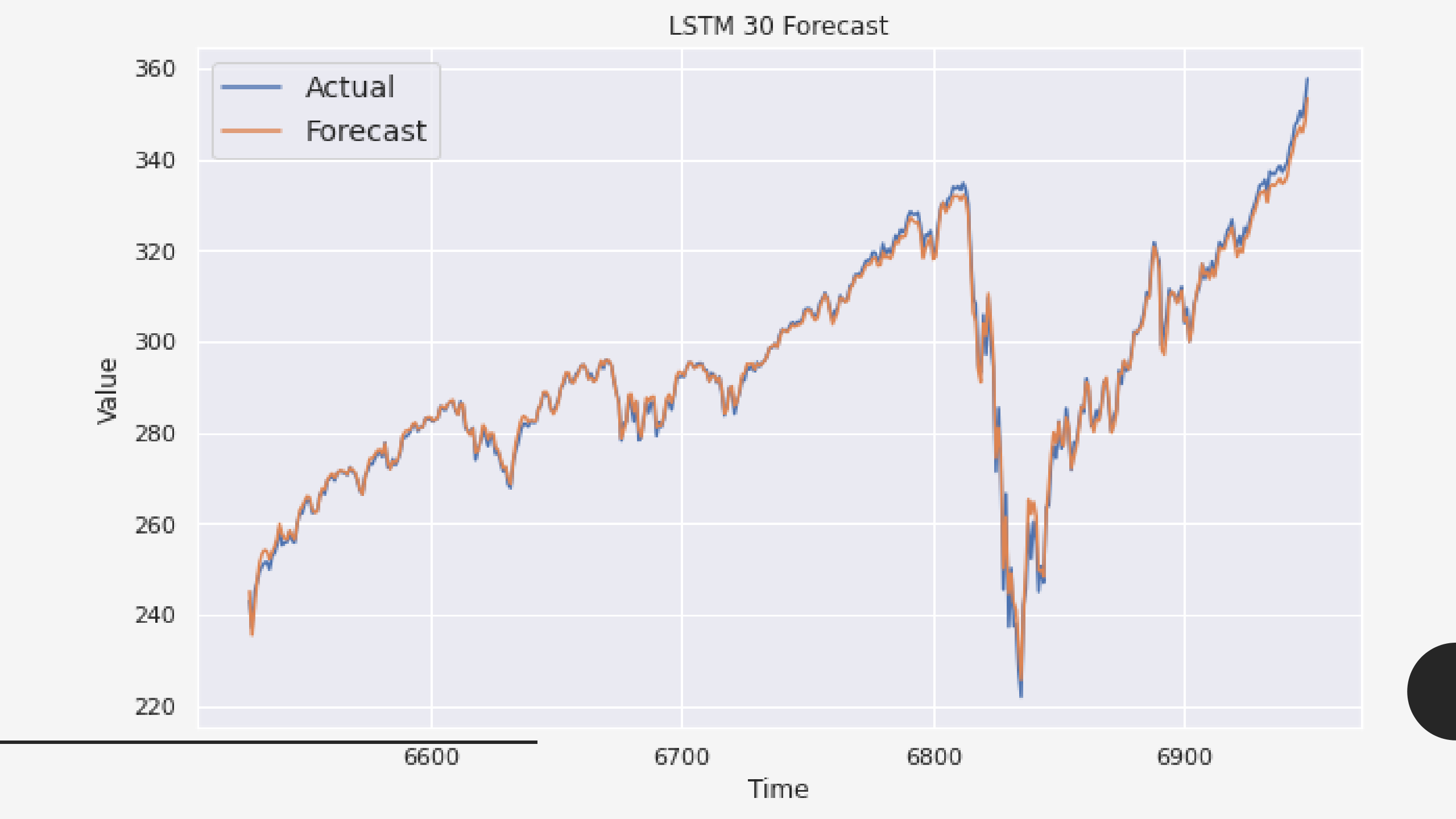
Long Short Term Memory

L Libraries

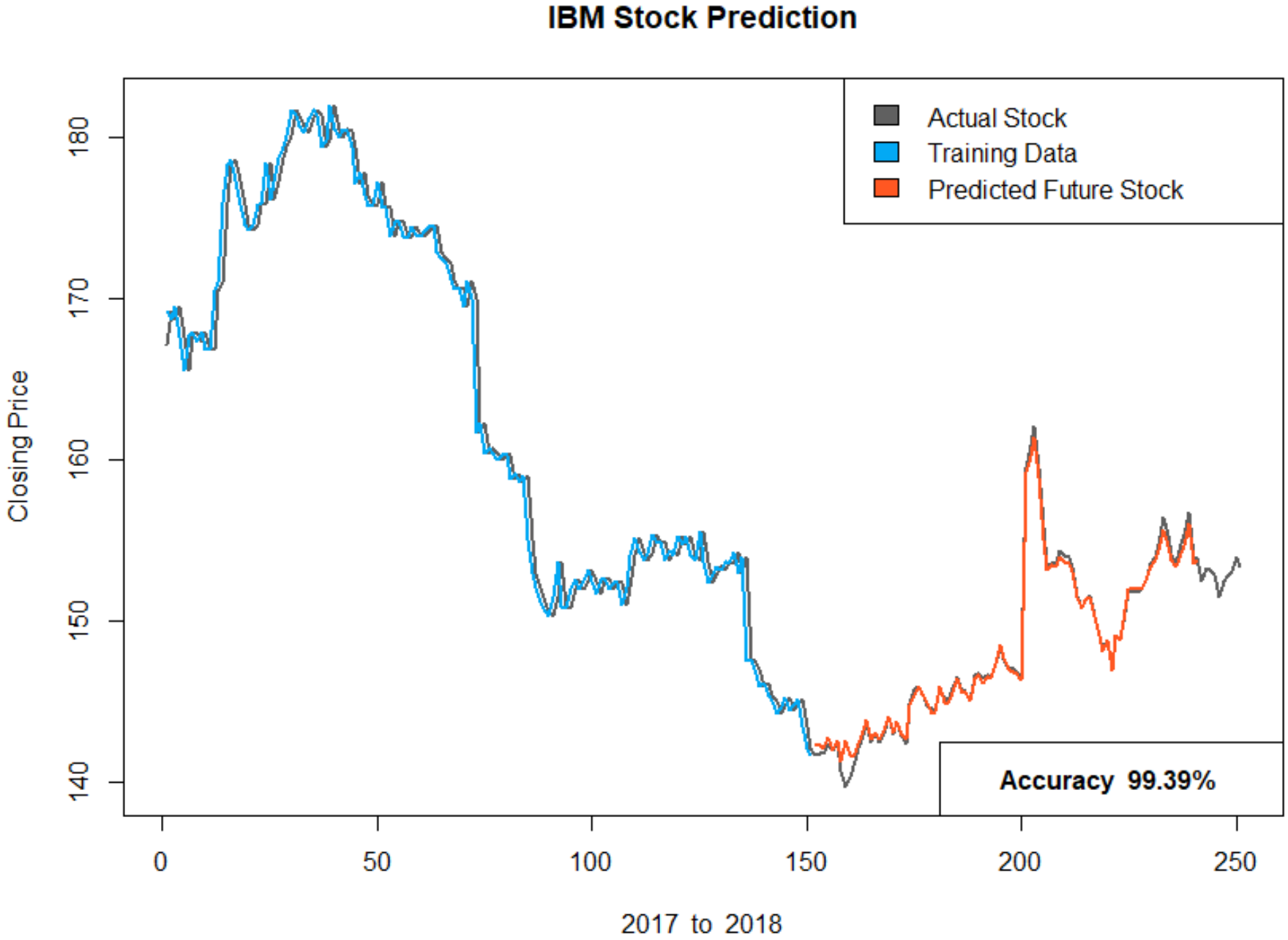
Pandas, Numpy, Sklearn, Tensorflow, etc

S Streamlit UI

Provided User Interface using Streamlit

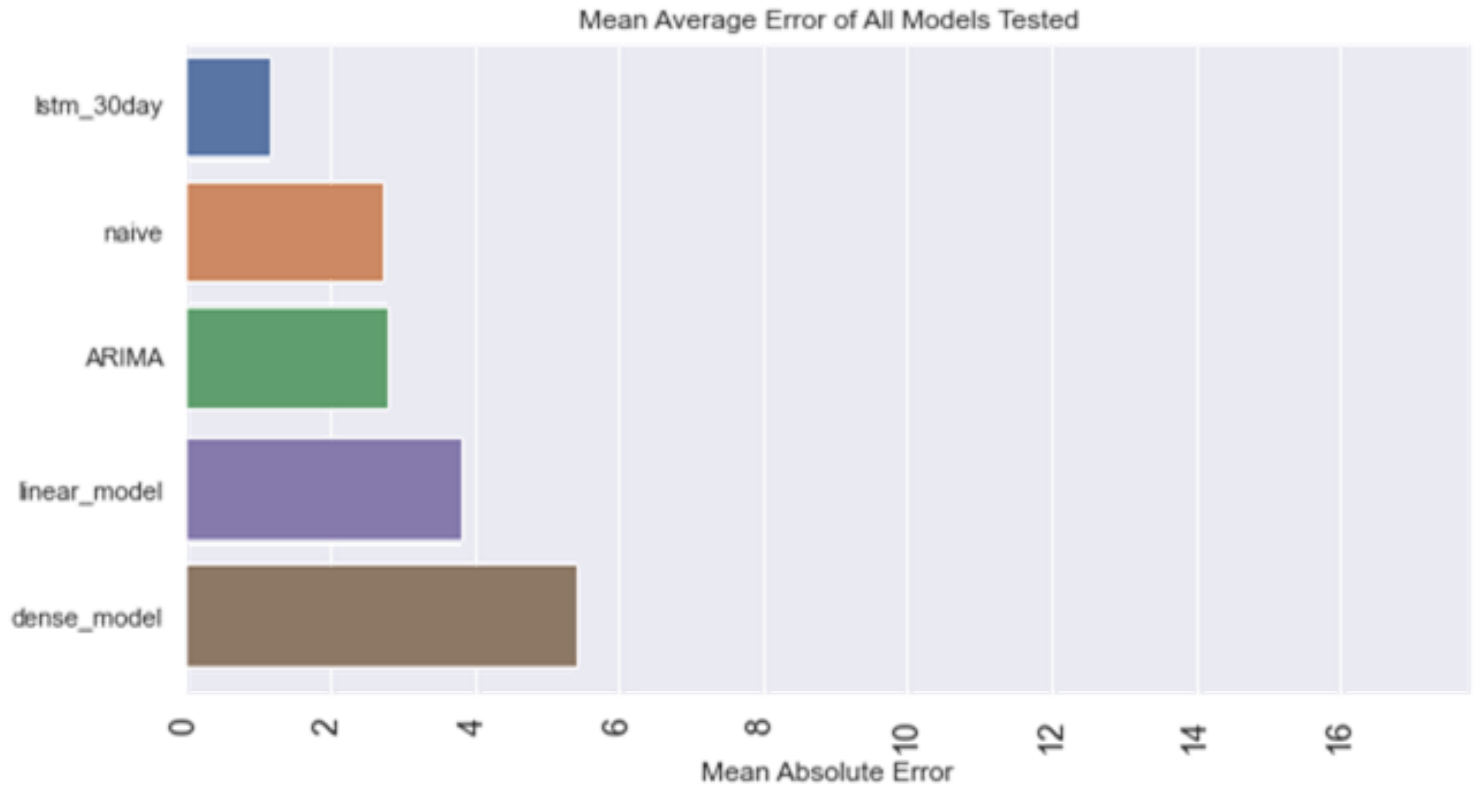

LSTM 30 Forecast
Actual
Forecast
360
340
320
300
280
260
240
220
Value
6600
6700
6800
6900
Time

Results & Discussions

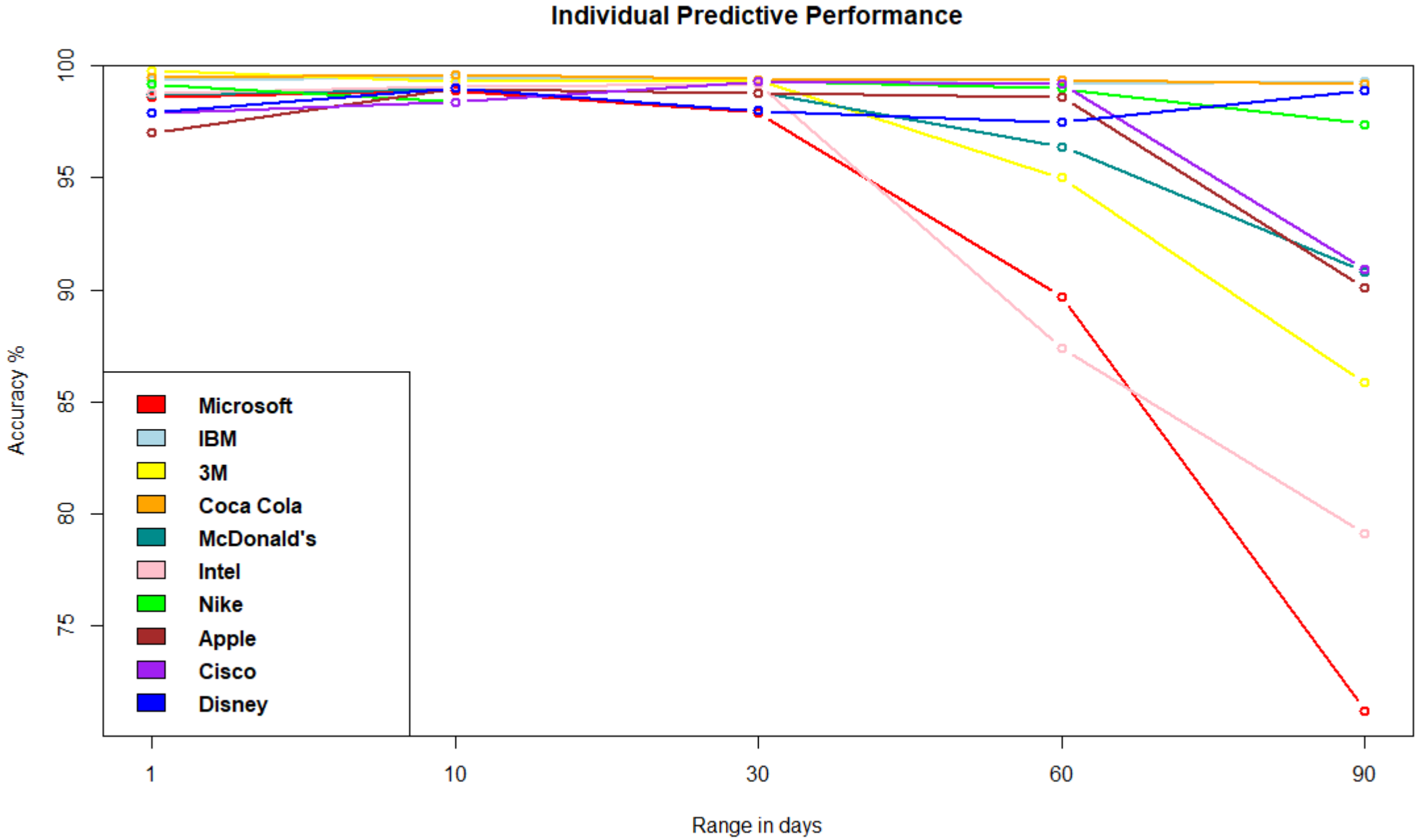


Project Summary

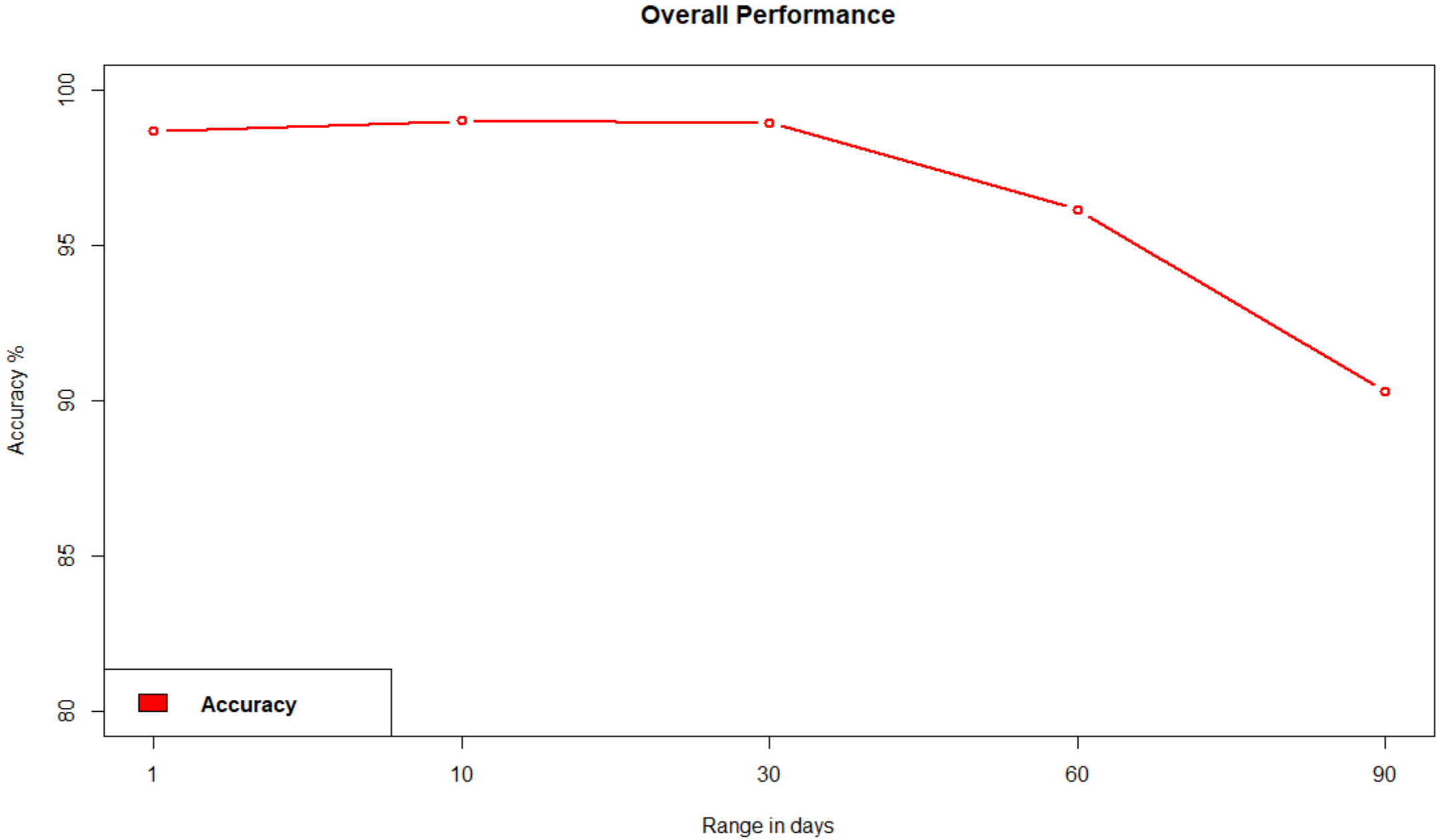
The resulting error from all models built and tested are shown in the bar graph below.



Performance



Performance



Select project

APPL

Select period

2011-02-12_2011-02-18

Select report

Data drift

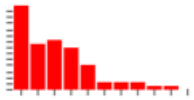
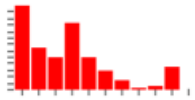
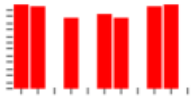
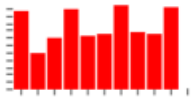
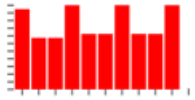
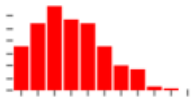
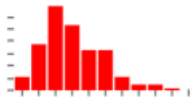
Project: APPL

Report:

- period: 2011-02-12_2011-02-18

APPL

Search

	Timeperiod	Current Status	Prediction	Prediction	Drift Score
>	4 Hour			Positive	0.019459
>	Hourly			Positive	1
>	15 minutes			Positive	1
>	1 minute			Negative	0.000001

What we Learnt

- Data Manipulation like cleaning of data, train-test-split, and model revaluation.
- Evaluation Metrics like mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE)
- Model Selection because choosing the right machine learning algorithm is essential.
- Deployment considerations for ease of access, distribution and usage.
- Limitations And Challenges because predicting stock prices accurately is challenging due to the complexity and volatility of financial markets.]

Applications in the Real World

- **Algorithmic Trading:** Automated trading decisions for profit in financial markets.
- **Portfolio Management:** Optimizing investment allocation and risk using price forecasts.
- **Risk Management:** Predicting losses and market trends to mitigate investment risks.
- **Risk Assessment for Lending:** Evaluating creditworthiness and default likelihood.
- **Hedging Strategies:** Protecting against adverse price movements.

Future Enhancement

- The limitation of the proposed system is its computational speed, especially with respect to sliding-window validation as the computational cost increases with the number of forward day predictions.
- The proposed model does not predict well for sudden changes in the trend of stock data.
- This occurs due to external factors and real-world changes affecting the stock market.
- We can overcome this by implementing Sentiment Analysis and Neural Networks to enhance the proposed model.
- We can modify the same system to an online-learning system that adapts in real-time.

Conclusion

Thus, as we can see above in our proposed method, we train the data using existing stock dataset that is available. We use this data to predict and forecast the stock price of n-days into the future.

The average performance of the model decreases with increase in number of days, due to unpredictable changes in trend.

The current system can update its training set as each day passes so as to detect newer trends and behave like an online-learning system that predicts stock in real-time.

Thank You