



Novel Approaches to Sentiment Analysis for Stock Prediction

AGRIM RAWAT
JEWEL BENNY
KNV KARTHIKEYA
MUDIT GAUR



TEAM 1



Introduction

- In the world of finance, stock market forecasts have been a crucial and contentious topic.
- Many people believe and think that it's nothing but a gamble and the stocks and their valuation can't be predicted.
- Through fundamental analysis, technical analysis, and, more recently, machine learning, others have tried to forecast the market.
- A technique like machine learning may be well suited for such an application given that the stock market is primarily a quantitative one.



Motivation and Aim

01

The current generation of machine learning models has concentrated on sentiment or technical analysis as a single feature.

02

However, as the stock market also largely relies on investor emotions and fundamental firm data, which cannot be represented by a straightforward numerical indicator.

03

Hence a model is proposed where the model takes in both stock financial data and news information, which we encode into a fixed-length vector.

04

We saw a lot of promise in the related work utilising sentiment to forecast stock movement. We were worried that a single mood value wouldn't adequately convey the nuanced nature of the news and corporate information, though.



Proposed Model(s)

- In order to more effectively process news data, a model encodes the data into real vectors before feeding the entire vector into a classification model.
- By merging the trade data for the stock and the data from the daily news that is relevant to the ticker, one may predict the stock's movement the next day.
- We divide our model into 2 subparts, one being representing the news from our sources as a fixed-length real scalar or vector and then combining this information with trading information, sentimental data, tech indicators etc.
- The stock movement prediction model needs the text representation model to transform news sentences and excerpts into real-value fixed-length scalar or vector representations.



Data Used for Prediction

- Headlines and Snippets related to the Stock market scrapped from Google news and other sources.
- Daily extraction of trading-related data, such as price and volume, from each stock ticker using the Yahoo Finance API.

Scraped News and Trading Data

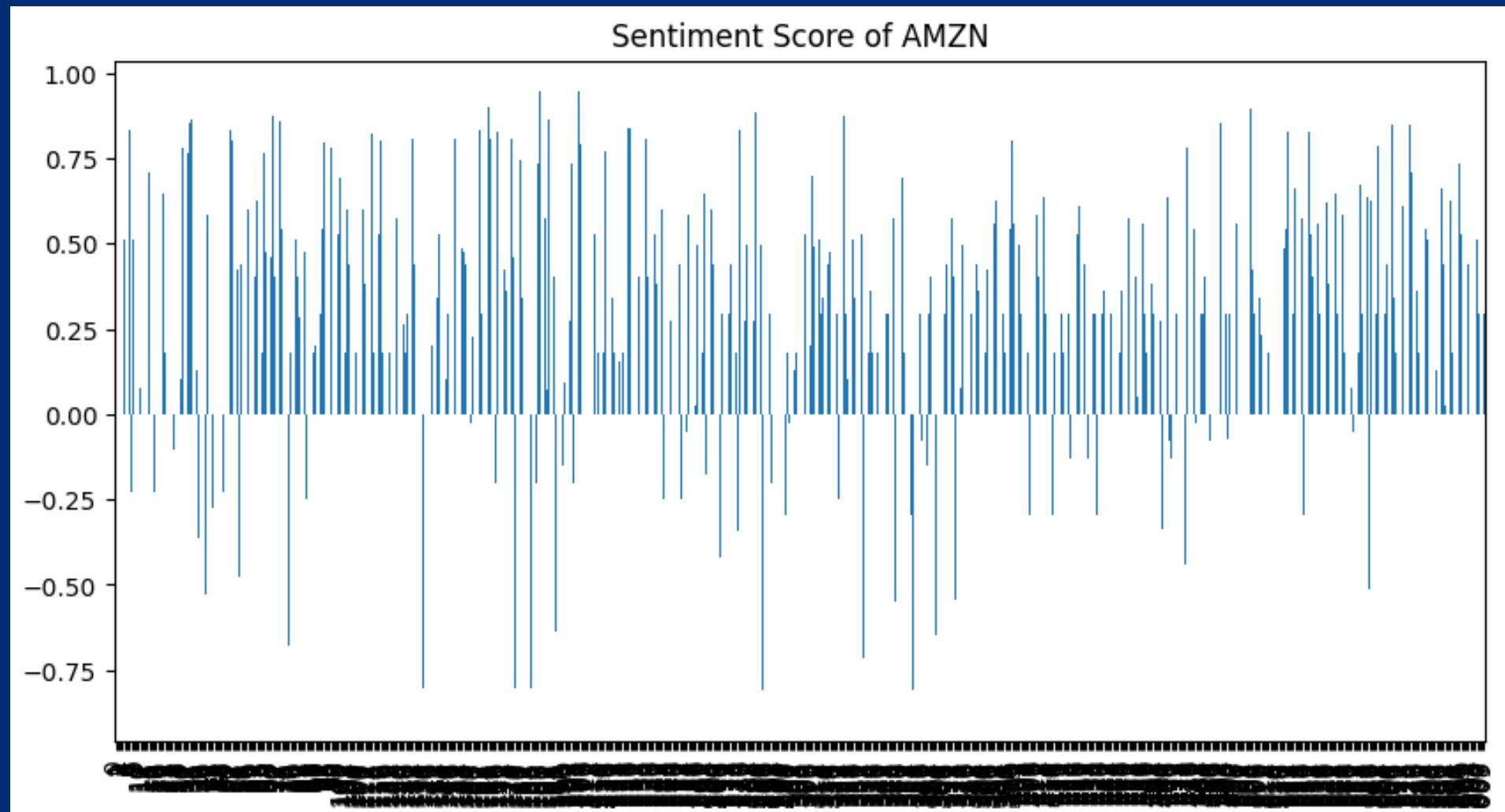
Date	Ticker	Url	headline	source	snippet
1 month ago	AAPL	https://www.apple.com	Apple celebrates the holiday	Apple	Celebrate the Holidays with Apple
3 days ago	AAPL	https://www.fiercehealthcare.com	Amwell exploring \$200M buy	Fierce Healthcare	Shares of online therapy company
1 week ago	AAPL	https://www.fiercehealthcare.com	State AGs call on Apple to say no	Fierce Healthcare	The Apple spokesperson also
1 day ago	AAPL	https://www.marketeconomy.com	Groupe Ldlic : H1 2022/2023	MarketScreen	An Apple reseller for businesses and
3 weeks ago	AAPL	https://finance.yahoo.com	YETI Holdings Q3 Earnings Tc	Yahoo Finance	The company held \$77.8 million in
3 weeks ago	AAPL	https://www.fool.com	This Tech Stock Is Down 40% The Motley Foc	The Motley Fool	The chipmaker, which is known for
01-Oct-22	AAPL	https://seekingalpha.com	Apple: Hello Recession (NASI	Seeking Alpha	Apple: Hello Recession Â· Summary Â·
23 hours ago	AAPL	https://thecyberwire.com	Heliconia described. Risks to CyberWire	CyberWire	Ukraine war latest news: Russian
01-Oct-22	AAPL	https://www.hellenic-shipping.com	STOCK MARKET SNAPSHOT FOR	Hellenic Shipping	STOCK MARKET SNAPSHOT FOR
2 weeks ago	AAPL	https://finance.yahoo.com	ASIA PODCAST FESTIVAL AND	Yahoo Finance	... to share information and build up
02-Oct-22	AAPL	https://seekingalpha.com	It's Not Over Until It's Over F	Seeking Alpha	It's Not Over Until It's Over For Apple,
28-Sep-22	AAPL	https://seekingalpha.com	Apple Stock: The Next Domir	Seeking Alpha	Apple shares lost as much as 4.5%
02-Oct-22	AAPL	https://seekingalpha.com	Apple: Understanding The Ba	Seeking Alpha	Shares of Apple have outperformed
30-Aug-22	AAPL	https://seekingalpha.com	Apple Stock: Buy Now And Ti	Seeking Alpha	Apple (NASDAQ:AAPL) stock has been
29-Sep-22	AAPL	https://seekingalpha.com	Apple Stock: Market Finally C	Seeking Alpha	-The news outlet noted that Blevins
02-Oct-22	AAPL	https://www.fool.com	Will Amazon Be Worth More The Motley Foc	The Motley Fool	But today Amazon is still worth \$1.2
29-Sep-22	AAPL	https://seekingalpha.com	Apple Services Is On A Critica	Seeking Alpha	The Apple stock's performance this
22-Sep-22	AAPL	https://seekingalpha.com	Apple Stock: Word Of Cautio	Seeking Alpha	Stocks go up or down depending on
30-Sep-22	AAPL	https://seekingalpha.com	Watch Apple and Tesla for a	Seeking Alpha	Krinsky noted that AAPL and TSLA,
03-Oct-22	AAPL	https://www.cnbc.com	Stocks surge to kick off	CNBC	The Dow shed 8.8% in September,
03-Oct-22	AAPL	https://seekingalpha.com	Apple Stock: Enough Is Enoug	Seeking Alpha	The other bit of good news is that this
03-Oct-22	AAPL	https://seekingalpha.com	Apple Stock: Current Price Is	Seeking Alpha	Is that the same Kevin O'Leary who
03-Oct-22	AAPL	https://seekingalpha.com	Keep Buying, Stocks Still Bott	Seeking Alpha	This is why the market is just not

Date	Open	High	Low	Close	Adj Cl...	Volume
1451	118	118.1699...	113.8799...	114.5599...	114.5599...	54678000
1452	115.0999...	116.25	112.4300...	113.6699...	113.6699...	42339700
1453	112.7099...	115.4800...	110.3899...	112.2099...	112.2099...	56432200
1454	112.4899...	113.8300...	111.4000...	112.9000...	112.9000...	45728700
1455	107.8799...	113.4400...	105.3499...	112.5299...	112.5299...	86868100
1456	114.0999...	114.9599...	106.5999...	106.9000...	106.9000...	67737300
1457	110.1100...	114.1900...	110.0899...	113.7900...	113.7900...	62782000
1458	119.0599...	119.5199...	114.7900...	116.3600...	116.3600...	65607400
1459	114.7099...	116.5899...	113.2200...	115.0699...	115.0699...	47198100
1460	113.8300...	118.2399...	113.5100...	115.25	115.25	48795100
1461	114.7900...	119.5899...	114.5	119.3199...	119.3199...	55660500
1462	119.9800...	120.3899...	116.5699...	119.8199...	119.8199...	49531500
1463	119.6500...	121.3199...	118.9499...	120.5999...	120.5999...	50934600
1464	116	119.3499...	114.7600...	115.6600...	115.6600...	68802300
1465	113.9199...	114.1200...	109.7699...	110.9599...	110.9599...	129605400

APPLE Scrapped News

AMAZON Trading Data

Sentiment Analyzer



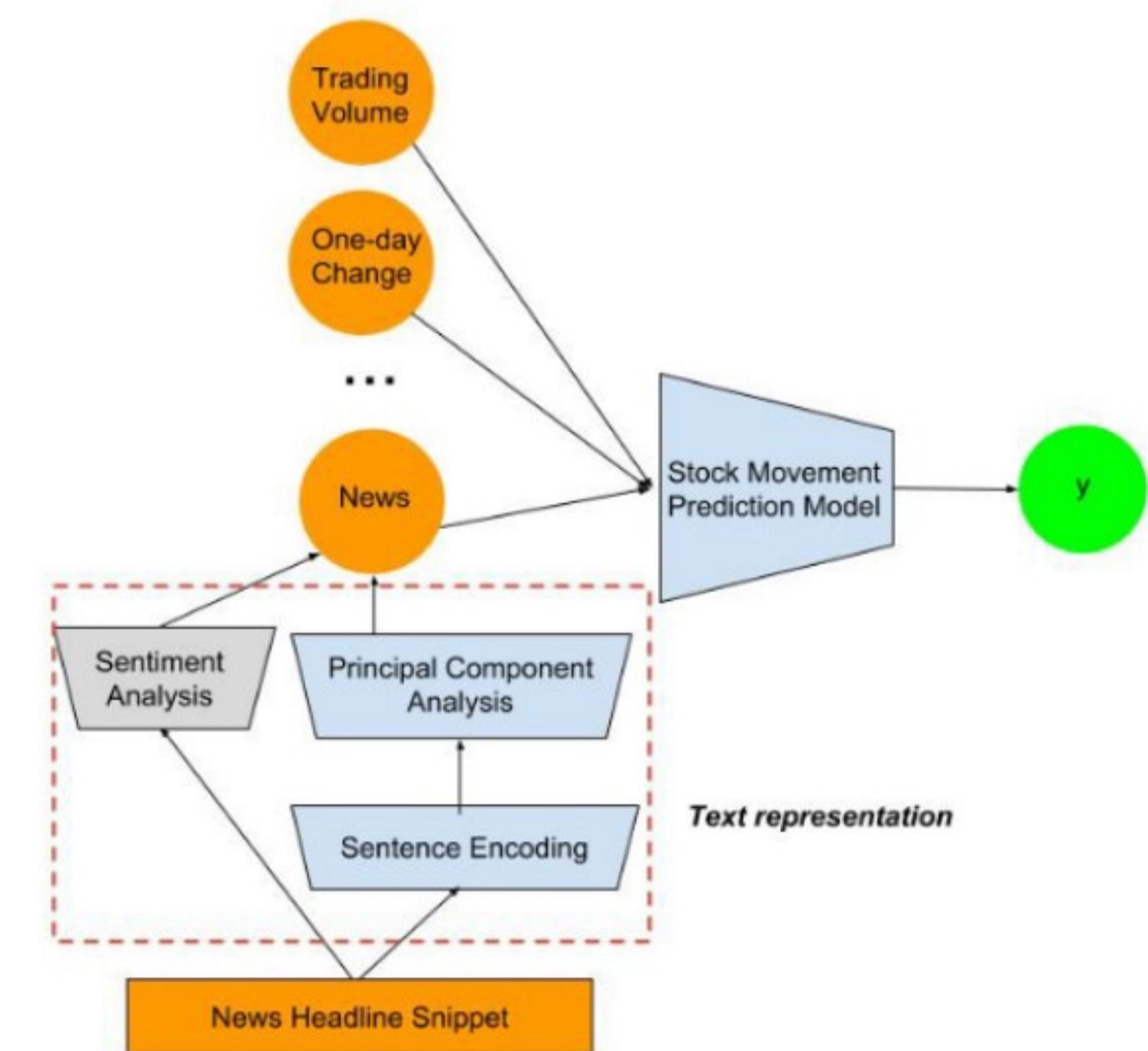
- Sentiment analysis is a process of analyzing the sentiment of peoples from various source.
- If the headline is postive for the company the sentiment score will be +ve and vice versa. It will be 0 for neutral news.
- For each headline, the compound score returns a normalized value between -1 (the most extreme negative headline) and 1 (the most extreme positive headline). Then, we create a daily score by averaging all the scores for the individual headline that we obtained that day.

**Sentiment Scores for News Snippets
about Amazon**



TEAM 1

The Model Design



Stock Movement Prediction and Various ML Models Used

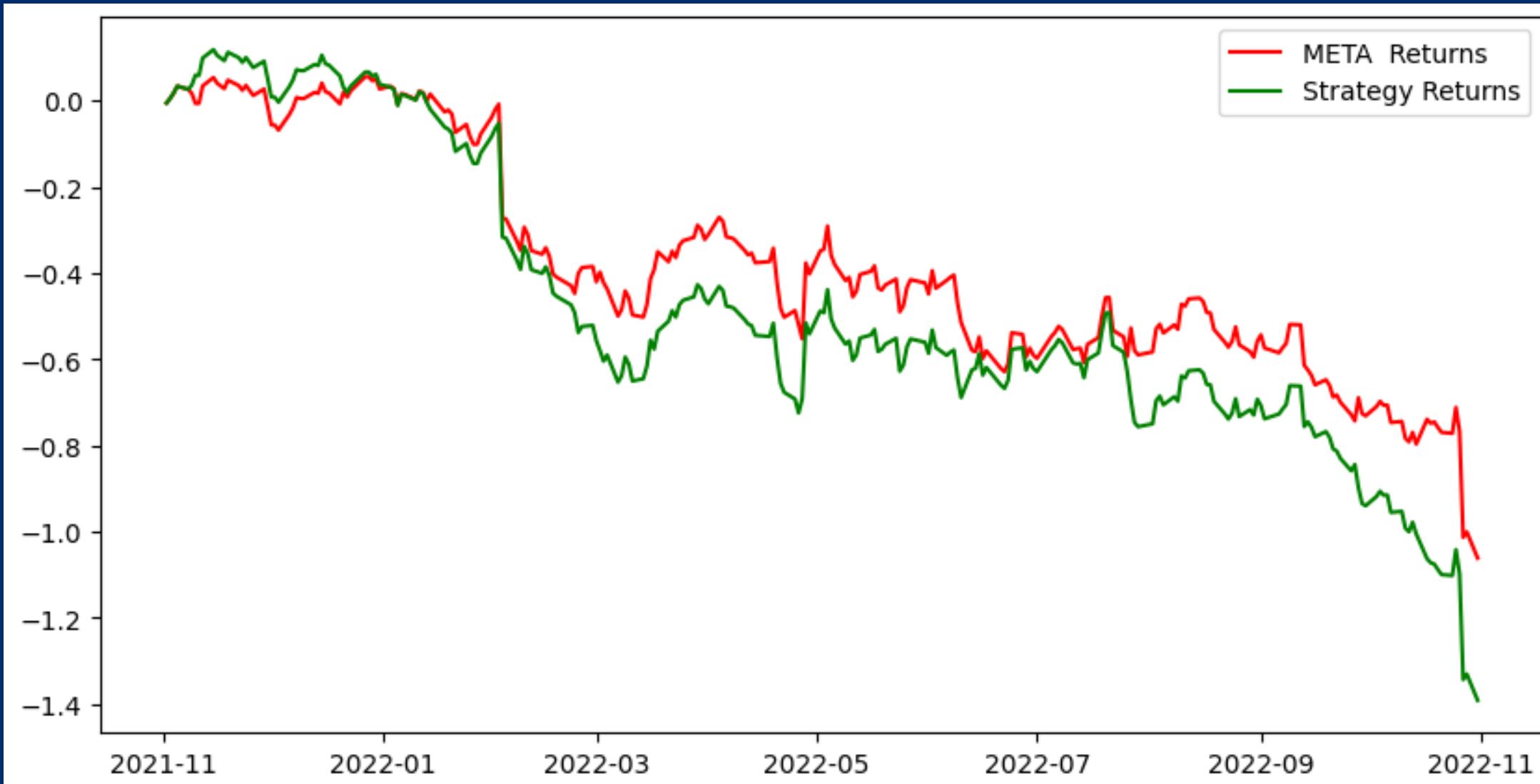
Random Forest is used for classification and regression, which builds a large number of decision trees during training time and produces the class that represents the majority vote of the individual leaves.

Support Vector Machines are mentioned in previous research to be effective in stock prediction applications. We can regulate the model by using different costs.

Bi-LSTM (Bidirectional Long Short Term Memory) neural network model is used along with sentiment data to make stock predictions keeping into account the expected behaviour from investors after seeing news headlines

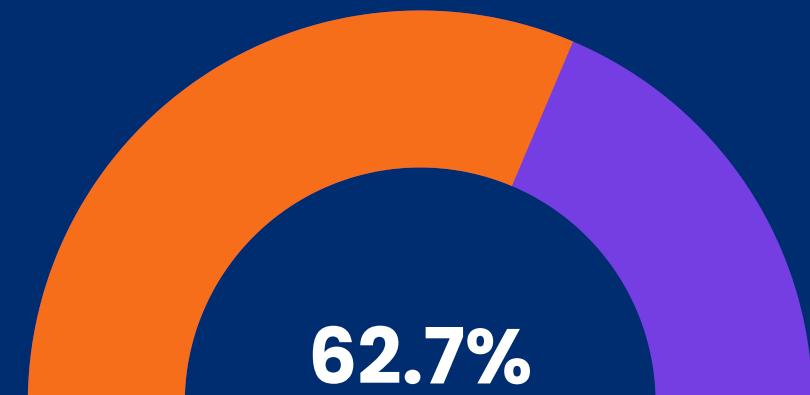
Recurrent Neural Networks or LSTMs are proven to be effective in dealing with sequential data, with the output being dependent on the previous computations. The hidden layer captures information about what has been calculated so far.

SVM

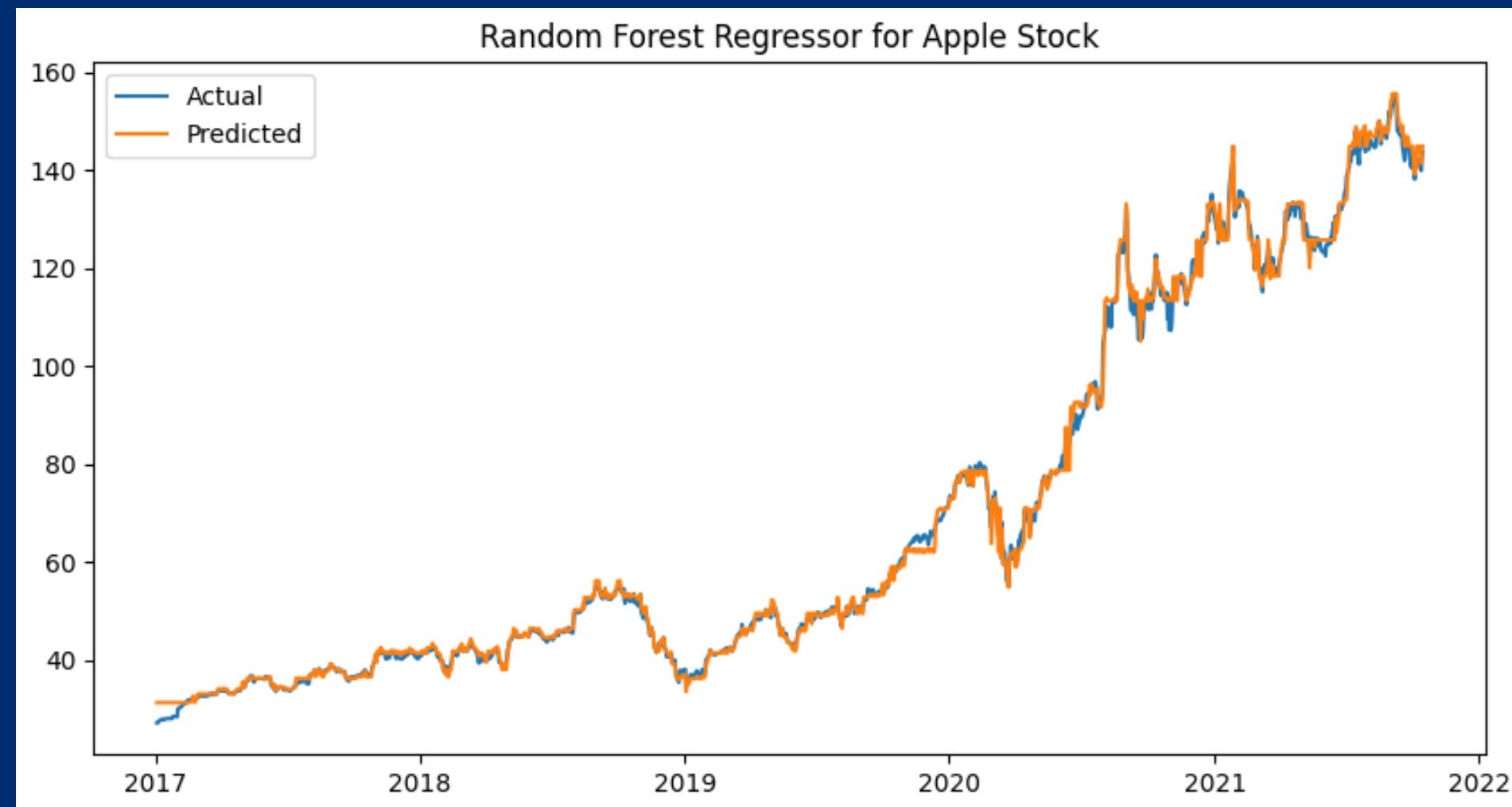


Ticker Used: META (Meta)
Accuracy = 62.745098%
Mean Absolute Error = 0.74509803

- The RBF kernel captures the high-dimensional nature of stock movement.
- We can regulate the model by using different costs C .
- The kernel transforms our input space into higher dimensional space where we can separate the classes.



Random Forest Regressor

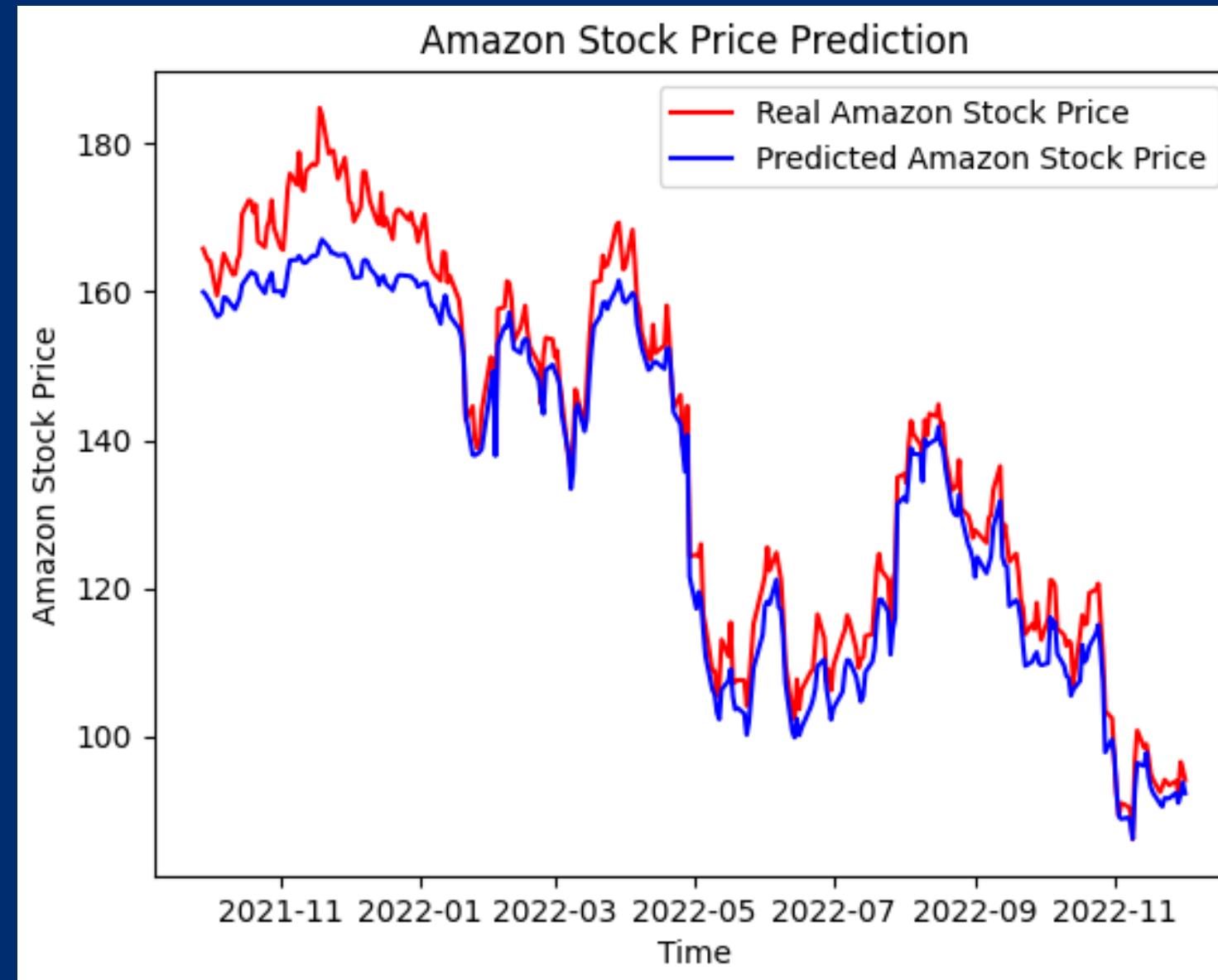


- Random Forests are used to fit the trading data against predicted stock prices.
- `RandomizedSearchCV()` is used to filter out the best parameters for the random forest regressor.

Ticker Used: AAPL (Apple)
Accuracy = 99.83%
Mean Absolute Error = 0.9844489768488489



Neural Network



Ticker Used: AMZN (Amazon).
Accuracy (with tolerance 10) = 92.97658%
Mean Absolute Error = 5.01572

- The neural network model used is based on LSTM (Long short-term memory).
- LSTMs are a type of RNN.
- Classic RNNs can keep track of arbitrary long-term dependencies in the input sequences.



Bi-LSTM



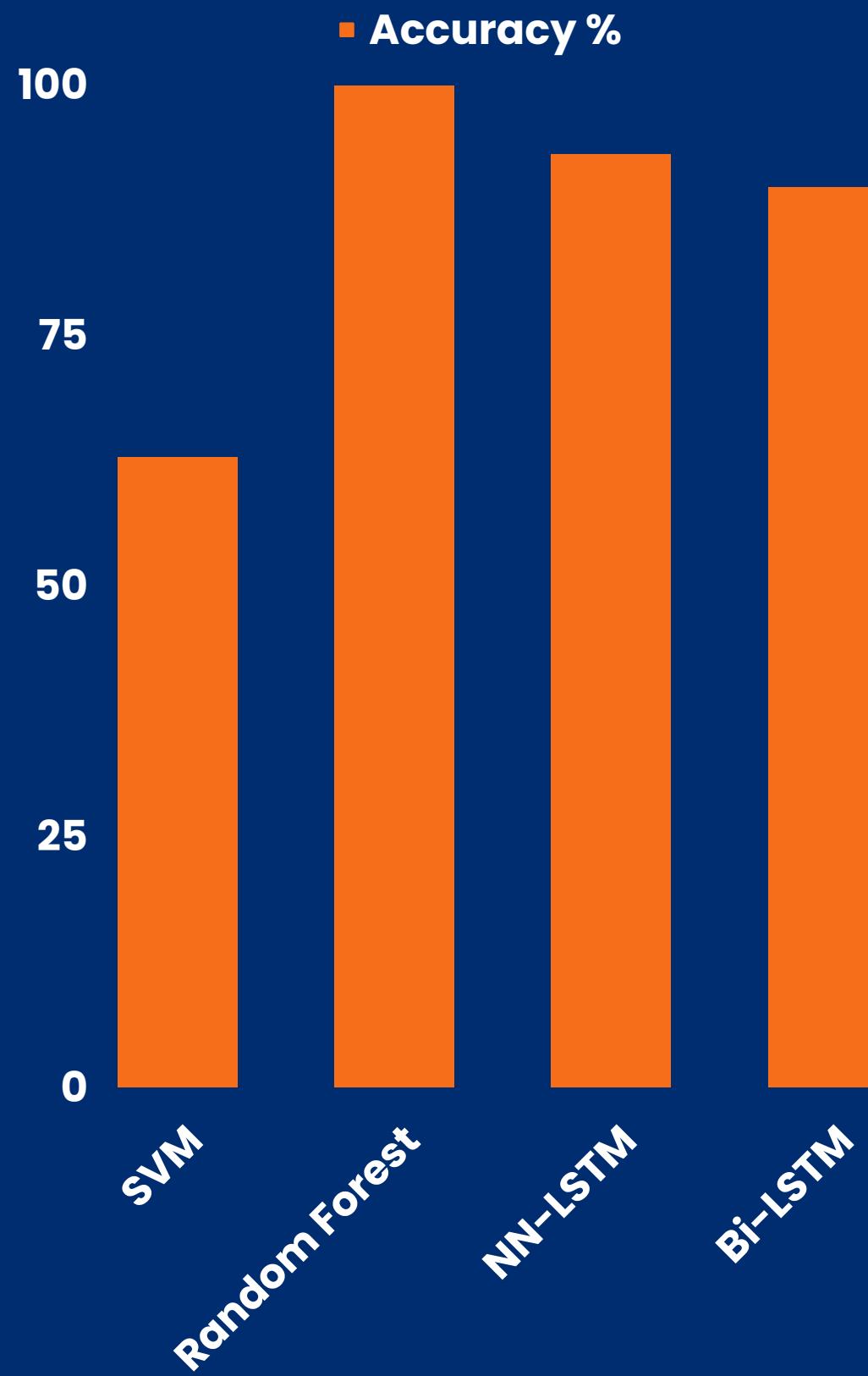
- Bidirectional - Long Short Term Memory, abbreviated to LSTM, is a segment of Recursive Neural Network (RNN) engineering (an artificial neural network)
- BiLSTM is capable of managing long time dependencies with the aid of using the usage of a memory unit. It has a sequence like structure, having 3 neural community layers or gates that are applied to the usage of diverse mathematical functions.

Accuracy : 89.69754253308129 %

Mean Absolute Error: 18.17



Results and Observations



- The stock predictions are very sensitive to the trading data provided.
- The random forest and NN-LSTM models seem to overfit the input data.
- The model for a specific company ticker is not transferable to a different company since the stock price behaviour changes from company to company.
- The sentiment-based Bi-LSTM seems to give better all round results even though the MAE is high.



Contributions

- **KNV Karthikeya:** Sentiment Analyzer, Bi-LSTM
- **Jewel:** News Web Scraping, NN-LSTM, Random Forest Regressor
- **Mudit:** Random Forest Regressor, SVM
- **Agrim:** SVM, Sentiment Analyzer



References

01

Novel Approaches to Sentiment Analysis for Stock Prediction
Chris Wang, Yilun Xu, Qingyang Wang – cs229stockprediction

02

Bollen et al. (2010) Twitter Mood Predicts the Stock Market. Journal of Computational Science.

03

Madge, S., Bhatt, S. (2015). Predicting Stock Price Direction using Support Vector Machines. Independent Work Report Spring

04

Beautiful Soup Documentation –
<https://www.crummy.com/software/BeautifulSoup/bs4/doc/#>



TEAM 1

Thank You