Stock Price & Volatility Analysis

01. Introduction

02. Brief discussion of related work

03. Problems addressed

04. Approach
- Classical Models
- Using Sentiment Analysis

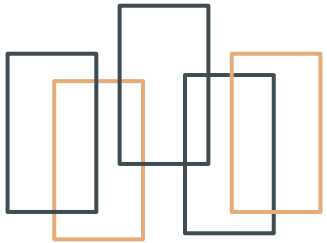
05. Data sets and experimental setup needed

06. Experiments we performed to prove our claim

07. List of Papers referred

08. Timeline

Introduction



Stock market prediction is an attempt to forecast the future trend of an individual stock, a particular sector or the market, or the market as a whole.

Approaches Considered-

- Classical Deep Learning Models
- Using Sentiment Analysis

Brief Discussion of Related Work

Neural networks for stock price prediction (2018)

Ren-Jie Han, Yu-Long Zhou, Yue-Gang Song

Five neural network models, namely, back propagation (BP) neural network, radial base function (RBF) neural network, general regression neural network (GRNN), support vector machine regression (SVMR), least square support vector machine regression, are surveyed and compared with predictive capacity (LS-SVMR). They conclude that the BP neural network outperforms the other four models.

Stock Prices Prediction using Deep Learning Models (2019)

Jialin Liu, Fei Chao, Yu-Chen Lin, and Chih-Min Lin

Deals with the challenges of using deep learning models for predicting stock prices, since there is a lot of noise and confusion in stock price-related details. In order to denoise the data, they use (1-D) residual convolutionary networks and in order to forecast the stock price, long-short term memory (LSTM) is then used.

Stock Trend Prediction using News Sentiment Analysis

Kalyani Joshi, Prof. Bharathi H.N., Prof. Jyothi Rao

Focuses on the famous theory of stock prediction i.e the Efficient Market Hypothesis. This project is about taking non-quantifiable data such as a company's financial news articles and forecasting the future market trend with the classification of news sentiment. They developed three distinct classification models. Observations show that in all forms of testing, RF and SVM perform well. Naive Bayes performs well, but not in contrast to the other two.

Sentiment Analysis of Twitter Data for Predicting Stock Market Movements

Venkata Sasank Pagolu, 2016

Highlights that social media nowadays are a true reflection of public perception and opinion on current affairs. Forecasts based on public opinions shared on Twitter. Two separate textual representations, Word2vec and N-gram, have been used in the present paper to examine public feelings in tweets.

Problem addressed

Classical Deep Learning based models

We try to contrast and compare the effectiveness of Classical Deep Learning based models to predict the stock prices for stocks of a specific company.Here we use the following models and their common variations in our thesis.

1. Vanilla RNN
2. Long Short Term Memory(LSTM) Model
3. Gated Recurrent Unit(GRU)
4. Convolutional Neural networks(CNN)

Volatility Prediction using Sentiment Analysis

According to EMH Theorem, Investors are not always rational while investing and are often dictated by Greed and Fear. Therefore, in order to capture the investor's mood, we are building a RNN(Recurrent neural network) which uses sentiment analysis along with historical stock price data to observe if it provides any improvement over the classical approach.

Classical Approach

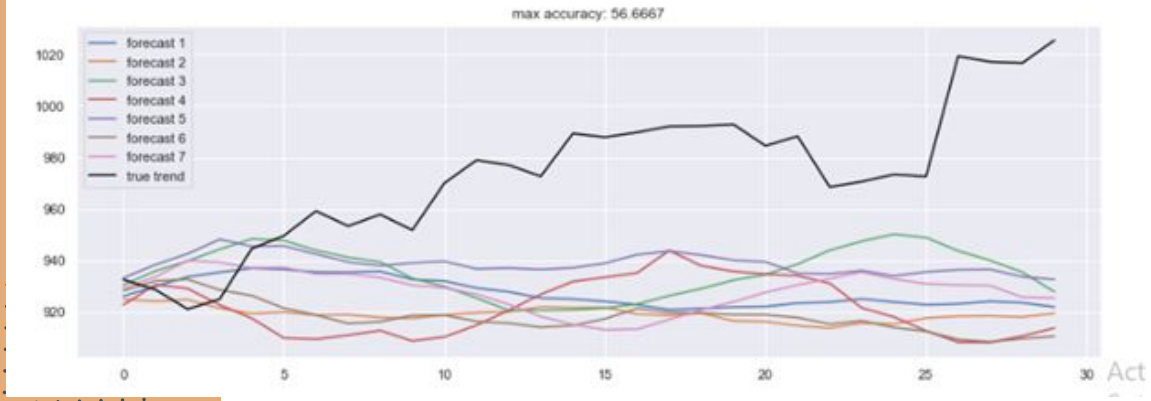
1. Historical Data Collection
2. Data Processing
    a. Normalisation
    b. Extraction of Close Stock Prices
    c. Train Test Split(220+30)
3. Models Considered-
    a. Vanilla RNN
    b. LSTM
        i. Vanilla
        ii. BI Directional
        iii. Sequence to Sequence
    c. GRU
        i. Vanilla
        ii. BI Directional
        iii. Sequence to Sequence
    d. CNN

Classical Approach Experimental Steps

1. Model Construction
    a. Parameters: num_layers = 1,size_layer = 128, timestamp = 5, epoch = 10, dropout_rate = 0.8,future_day = test_size, learning_rate = 0.01, test_size=30
2. Training
    a. 220 Days
    b. Batch Size=5 Data Points
3. Testing
    a. Test Size=30 Days
4. Running Simulations
    a. #Simulations=7
5. Calculating accuracy
6. Plotting Graphs

Predicted Results

LSTM



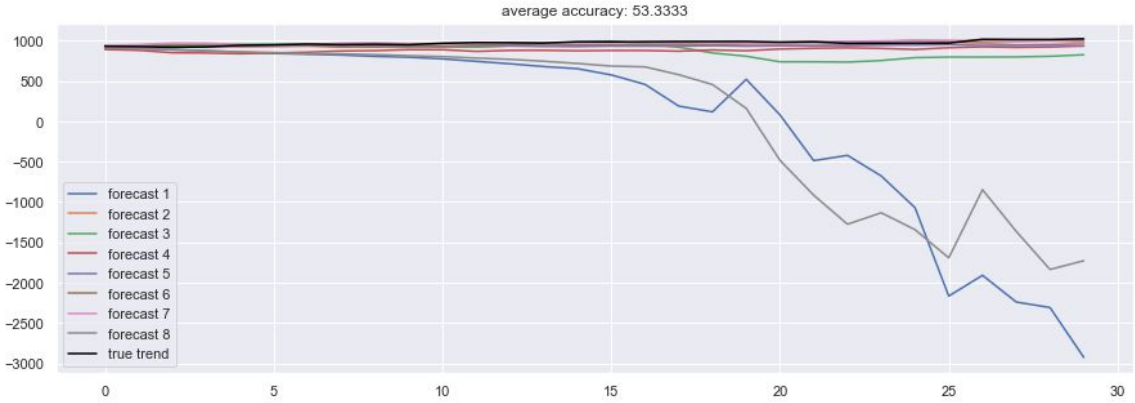
volatility accuracy=56.67%

Price accuracy = 94.446%

GRU

volatility accuracy=53.33%

Price accuracy = 96.227%



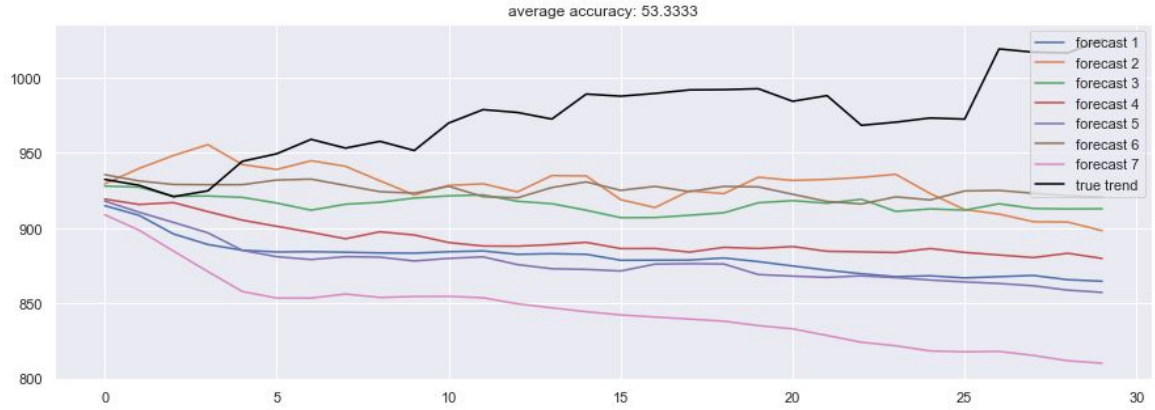
Predicted Results

RNN

volatility accuracy=53.3337%

Price accuracy = 91.4686%

CNN



volatility accuracy=55.2346%

Price accuracy = 90.7365%

Stock volatility prediction using Sentiment Analysis

News Article Generation

Web Scraper made in Python is used to get the list of links of news articles using “scrapy” module. We use “Reuters.com” website for this purpose.The role of this file is to scrap out a list of news articles from the website about a specific company mentioned by date

Sentiment Score Generation

We use “Rossete” API for text analysis for the contents for each article. This file outputs a csv file which contains the columns date, label, confidence, entity-label and entity-confidence

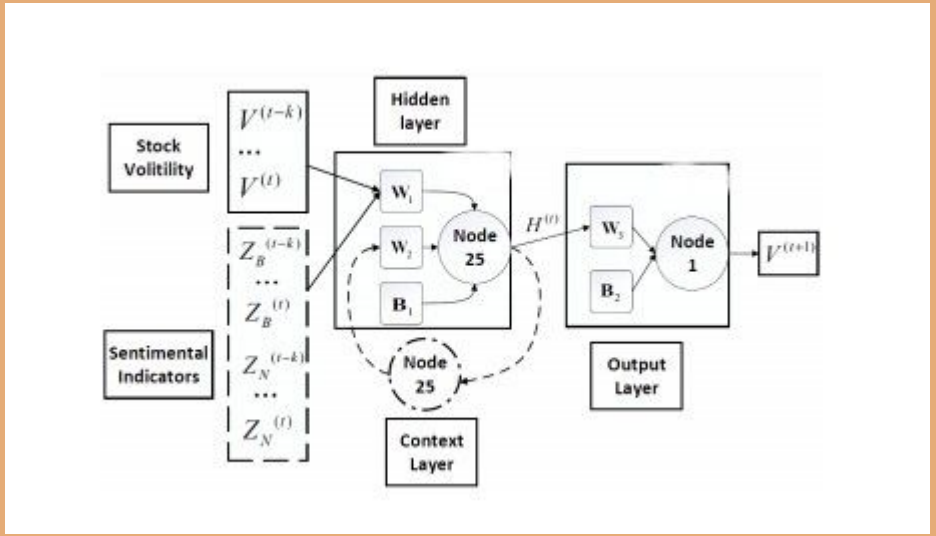
Calculating Bullishness And combine with Prices

We calculate the bullishness score of the selected stock on a per day basis. Bullishness for each day is calculated as: **Bullishness score** = (Sum of positive confidence – sum of Negative confidence)/(total articles)

Combine the sentiment analysis of that day with the corresponding stock price

RNN Model and Datasets

For training our model, the sentiment score goes z score normalization followed by min-max normalization. Then, we construct a 2 layer LSTM followed by 2 Dense layers. Input of which will be previous k days stock price and it's corresponding sentiment score where k is the window size. The output is the prediction of stock price. Hyperparameters of this model are as follows batch_size=128, epochs=200, validation_split=0, verbose=0).

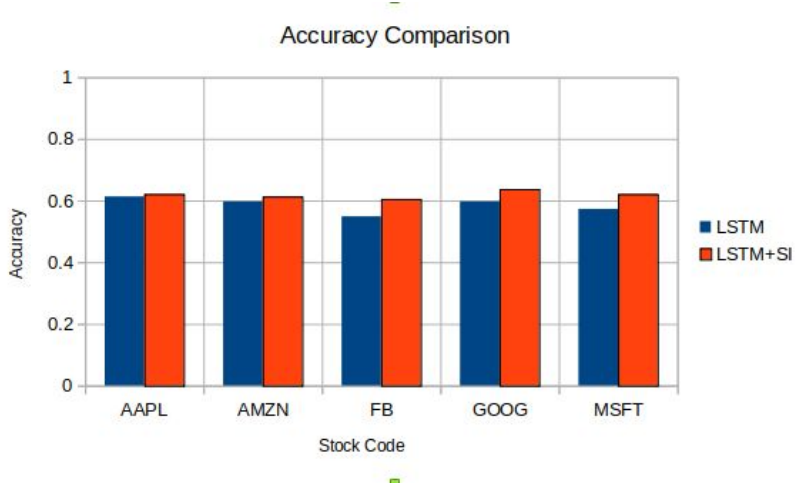


Datasets Used

- News articles from Reuters.com on particular stock ranging from 2012 - 2018
- Historical Stock Prices from Yahoofinance.com

Result and Conclusion

Accuracy and the best k for RNN+EMM and RNN					
	AAPL	AMZN	FB	GOOG	MSFT
LSTM	0.6129	0.5968	0.5484	0.5968	0.5726
LSTM+SI	0.621	0.6129	0.6048	0.6371	0.621
k(LSTM)	4	7	6	7	6
k(LSTM+SI)	4	9	7	7	5

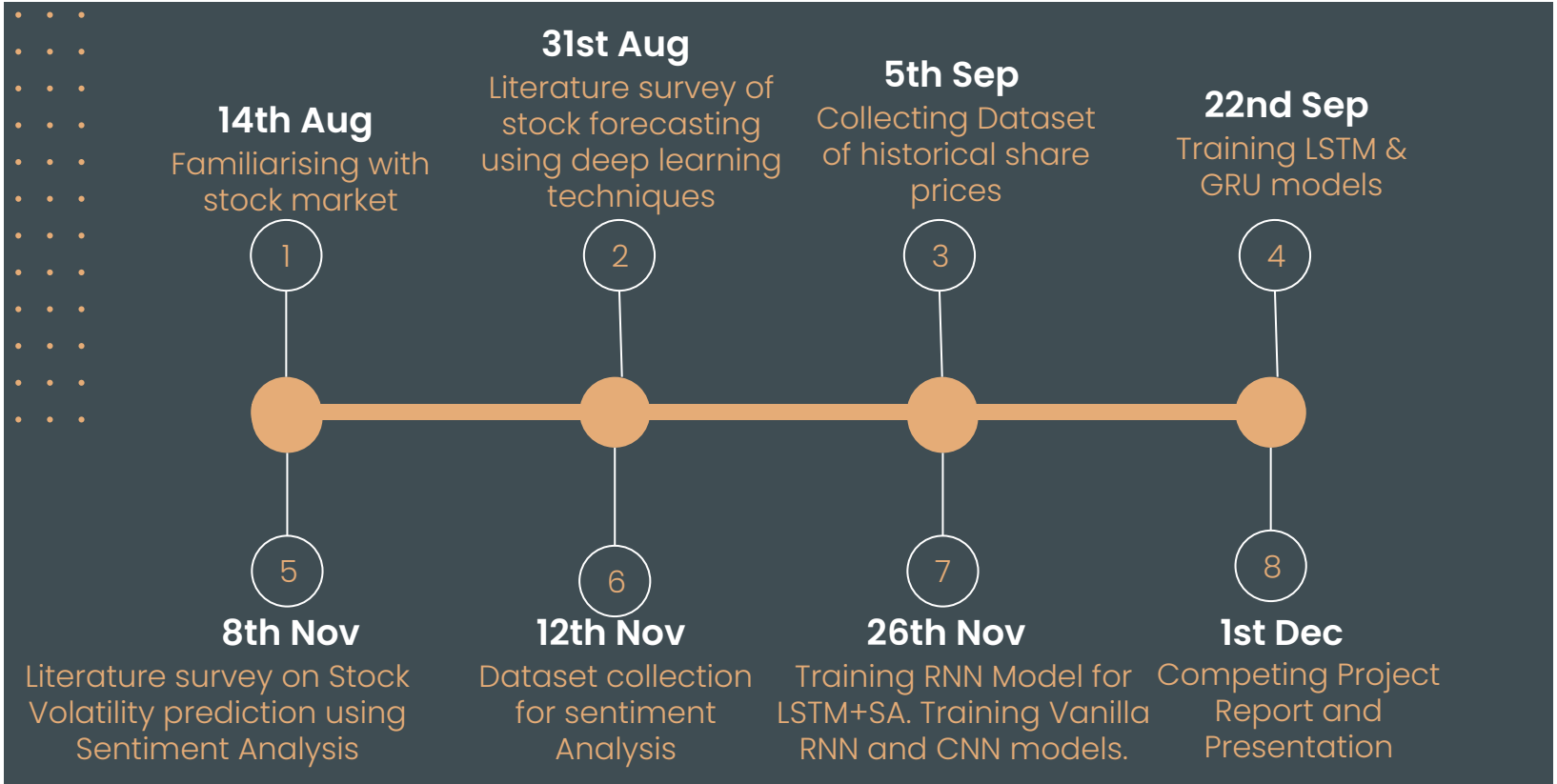


After completing the prediction, we can clearly observe that the results using sentiment analysis offer marginally better results as compared to predicting without them. Hence, we conclude that analyzing sentiments in order to predict the volatility in the stock market is not worth the effort.

List of Papers Referred

S. No	Research Paper	Author(s)
1.	Stock Prices Prediction using Deep Learning Models (2019)	*Jialin Liu, Fei Chao, Yu-Chen Lin and Chih-Min Lin*
2.	Neural networks for stock price prediction (2018)	*Ren-Jie Han, Yu-Long Zhou and Yue-Gang Song*
3.	Stock Trend Prediction using News Sentiment Analysis	*Kalyani Joshi , Prof. Bharathi H. N. and Prof. Jyothi Rao*
4.	Sentiment Analysis of Twitter Data for Predicting Stock Market Movements (2016)	*Venkata Sasank Pagolu*
5.	Stock Volatility Prediction Using Recurrent Neural Networks with Sentiment Analysis	*Yifan Liu, Zengchang Qin, Pengyu Li and Tao Wan*
6.	Market Trend Prediction using Sentiment Analysis: Lessons Learned and Paths Forward	*Andrius Mudinas and Dell Zhang*
7.	Stock Forecasting using M-Band Wavelet-Based SVR and RNN-LSTMs Models	*Xiaodi Wang and Abdul Hasib Rahimyar*
8.	A Robust Predictive Model for Stock Price Prediction Using Deep Learning and Natural Language Processing	*Sidra Mehtab and Jaydip Sen*

Timeline


14th Aug
Familiarising with stock market
1
31st Aug
Literature survey of stock forecasting using deep learning techniques
2
5th Sep
Collecting Dataset of historical share prices
3
22nd Sep
Training LSTM & GRU models
4
5
8th Nov
Literature survey on Stock Volatility prediction using Sentiment Analysis
6
12th Nov
Dataset collection for sentiment Analysis
7
26th Nov
Training RNN Model for LSTM+SA. Training Vanilla RNN and CNN models.
8
1st Dec
Competing Project Report and Presentation

Thanks!