

# Combining Time Series and Sentiment Analysis for TCS Forecasting

1<sup>st</sup> Kiran Khanna

*Dept. Computer Science With AI Engineering (AIE)*  
Amrita Vishwa Vidyapeetham  
AM.EN.U4AIE19037

3<sup>rd</sup> Dishith Reddy

*Dept. Computer Science With AI Engineering (AIE)*  
Amrita Vishwa Vidyapeetham  
AM.EN.U4AIE19041

4<sup>th</sup> Chaitanya Kapoor

*Dept. Computer Science With AI Engineering (AIE)*  
Amrita Vishwa Vidyapeetham  
AM.EN.U4AIE19022

**Abstract**—The goal of this project is to create a model that can forecast stock price utilising sentimental information from news headlines and historical prices, and the model must not only provide better outcomes but also reduce the disparity between predicted and real values. The influence of news headlines on stock prices may be seen. We have predicted stock price using different machine learning and deep learning models

**Index Terms**—Long short-term memory , Machine Learning , Stock market , Sentiment analysis

## I. INTRODUCTION

All Stock prediction is a difficult topic to solve in the fields of economics, engineering, and mathematics. It has drawn a lot of interest from both academics and business because of its financial gain. The majority of investors and financial professionals have long been fascinated by stock price forecasting. Nonetheless, because there are so many other factors that can impact stock prices, determining the ideal moment to buy or sell has remained a challenging assignment for investors.

Many things influence the direction of a trend. News and historical pricing are two of the most essential elements influencing the trend of a financial asset. Sentiment is a subjective attitude toward a subject, or an opinion used to portray emotion. Text mining, also known as sentiment analysis, may extract subjectivity from reviews, customer comments, or texts. Sentiment analysis is used to obtain important information in a variety of industries, particularly business. The feelings of customer reviews, for example, may be used to adjust corporate strategy, which is why sentiment analysis is so popular. Because historical stock prices are time series data, several time series methods are used to forecast stock prices. For stock price prediction they are machine learning and deep learning methods are used.

## II. LITERATURE REVIEW

[1]It has never been easy to invest in a group of investments; the abnormalities of the financial market prevent simple models from accurately predicting future asset prices.

Machine learning, which involves teaching computers to execute activities that would ordinarily need human intellect, is now the most popular study topic. This article tries to develop a model for predicting future stock market values using Recurrent Neural Networks (RNN) and, in particular, the Long-Short Term Memory model. The testing result conform that lstm model is capable of tracing the evolution of opening prices for both assets.

[2]The ARIMA model is used in this paper to build a comprehensive stock price prediction model. Stock price predictions are made using published stock data from the New York Stock Exchange (NYSE) and the Nigeria Stock Exchange (NSE). The results showed that the ARIMA model has a lot of potential for short-term prediction and can compete with existing stock price prediction strategies. This could guide investors in stock market to make profitable investment decisions. With the results obtained ARIMA models can compete reasonably well with emerging forecasting techniques in short-term prediction.

[3]The suggested approach is comprehensive since it incorporates stock market dataset pre-processing, several feature engineering techniques, and a custom deep learning-based system for stock market price trend prediction. We conducted extensive assessments on commonly used machine learning models and found that our suggested approach outperforms owing to the extensive feature engineering that we implemented. The technology predicts stock market trends with a high degree of accuracy. This paper adds to the stock analysis research community in both the financial and technological areas by providing extensive design and assessment of prediction term lengths, feature engineering, and data pre-processing approaches.

[4]In practise, time series forecasting entails building models that generalise data from previous values and make predictions for the future. Furthermore, it is reasonable to infer that financial time series forecasting incorporates phenomena formed in part by the social context. As a result, the current research focuses on the application of sentiment analysis methods to data gathered from social networks, as well as their

usage in multivariate prediction systems involving financial data. 22 distinct input configurations employing such extracted information were examined across a total of 16 different datasets, under the frameworks of 27 different algorithms, in an extended testing procedure. Two case studies were used to organise the comparisons.

### III. METHODOLOGY

#### A. Dataset

We have done sentiment analysis and time series prediction

1) *Sentiment Analysis*: We have extracted our dataset from FINVIZ.com for TCS stock news sentiment analysis.

	Ticker	Date	Time	Headline
0	TCS	May-03-22	04:05PM	The Container Store Group, Inc. Announces Four...
1	TCS	Apr-12-22	12:12PM	The Container Store Group, Inc.'s (NYSE:TCS) S...
2	TCS	Apr-11-22	08:18AM	Tata Consultancy Earnings Trail Estimates Afte...
3	TCS	Apr-05-22	02:45AM	Payments Canada Partners with TCS to Evolve Pa...
4	TCS	Mar-31-22	08:30AM	The Container Store Introduces New Loyalty Pro...

Fig. 1. TCS news headline dataset

2) *Stock market prediction*: We have taken our dataset from yahoo finance for TCS.NS stock market prediction.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2019-01-01	1896.000000	1910.000000	1885.000000	1902.800049	1785.039795	1094883
1	2019-01-02	1905.000000	1934.449951	1900.000000	1923.300049	1804.270874	2100463
2	2019-01-03	1919.000000	1944.949951	1893.099976	1899.949951	1782.366089	2611668
3	2019-01-04	1900.000000	1901.199951	1841.000000	1876.849976	1760.695679	4280862
4	2019-01-07	1891.800049	1908.800049	1881.000000	1897.900024	1780.442993	1856423

Fig. 2. TCS.NS stock price dataset

#### B. System Architecture

Taking sentiment from a news headline we will get to know whether a particular news headline to postive or negative for the stock market. By doing stock closing price prediction using different machine and deep learning algorithms , we can predict the closing pricing of tomorrow. By taking sentiment value and tomorrow closing price value we will come to know where the stock will rise or fall . We have also made a website , which will take a input as a stock name and does sentiment and stock closing price prediction and combine the both results to recommend whether the stock will rise or fall and should we buy or sell . In this project for sentimental analysis we have done various preprocessing techniques and applied Textblob to find the polarity of the text.

1) *NLP Text Preprocessing*: Text preprocessing is a technique for cleaning text data and preparing it for use in a model. **Lower case** Because the machine treats lower case and upper case differently, it is simple for a computer to read the words if the text is in the same case.

**Remove punctuations** There are a total of 32 primary punctuations that must be addressed. We may use a regular

expression and the string module to replace any punctuation in text with an empty string.

**Remove Stopwords** Stopwords are the most frequently occurring words in a text that offer no useful information. Stopwords such as they, there, this, where, and others are examples of stopwords. The NLTK library is a widely used library for removing stopwords, with around 180 stopwords removed.

**Stemming and Lemmatization** Stemming is the process of reducing a word to its basic stem, such as run, running, runs, and runed, which are all derived from the same word. The words are stemmed using the NLTK package. The stemming technique is not utilised in production since it is inefficient and frequently stems undesired words. As a result, another approach known as lemmatization was introduced to the market to overcome the problem. Porter stemmer and snowball stemmer are two examples of stemming algorithms.

2) *Textblob*: As a Lexicon-based sentiment analyzer, TextBlob It contains certain predetermined rules, or a word and weight dictionary, with some scores that assist compute the polarity of a statement.If the polarity is greater than zero then sentiment is positive , less than zero then sentiment is negative . if polarity is equal to zero then it is neutral sentiment .

For stock market prediction we have used lstm , linear regression , random forest,arima and prophet .

3) *LSTM*: Long Short-Term Memory (LSTM) is one of several forms of Recurrent Neural Network RNN that can capture data from previous stages and utilise it to make predictions in the future. Every LSTM node must have a set

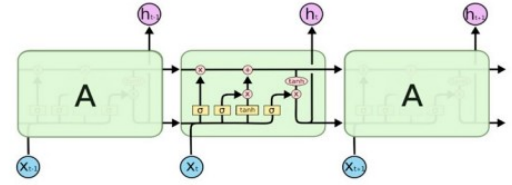


Fig. 3. LSTM

of cells responsible for storing passed data streams; the upper line in each cell connects the models as a transport line passing data from the past to the present; the independency of cells aids the model's dispose filter of adding values from one cell to another; and the independency of cells aids the model's dispose filter of adding values from one cell to another. Finally, by discarding or allowing input to pass through, the sigmoidal neural network layer constituting the gates drives the cell to an ideal value. Each sigmoid layer has a binary value (0 or 1), with 0 allowing no passage and 1 allowing all passage.

### IV. RESULTS AND DISCUSSION

For stock market prediction lstm and rlinear regression we got mse as 2.45 and 2.72 Website for these project we have build using streamlit and python libraries.



Fig. 4. Website sentiment analysis's

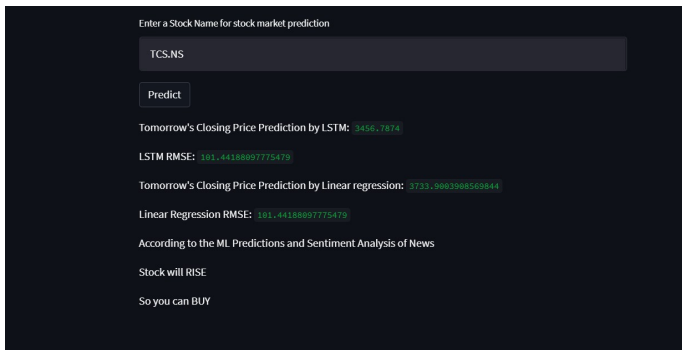


Fig. 5. Website stock market prediction

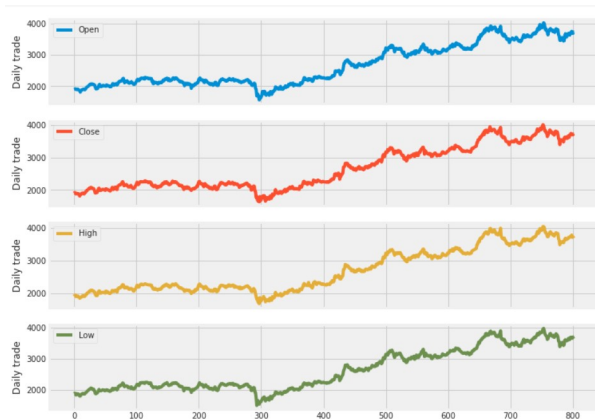


Fig. 6. Stock market

## V. CONCLUSION

Unlike traditional stock market prediction algorithms, our revolutionary technique integrates common people's feelings with news feed data and stock price to forecast stock market behaviour. Finviz provides a news feed of stock-related news, and sentiment polarity of news sentences is calculated to predict whether stock news is positive, negative, or neutral. Lstm and linear regression provide good mean square error for stock market prediction. Short data works well with Arima.

## VI. REFERENCE

### REFERENCES

- [1] Adil MOGHAR and Mhamed HAMICHE, Stock Market Prediction Using LSTM Recurrent Neural Network, International Workshop on Statistical Methods and Artificial Intelligence (IWSMAI 2020), University Abdelmalek Essaadi, Morocco, April 6-9, 2020.
- [2] yodele A. Adebiyi, Aderemi O. Adewumi and Charles K. Ayo, Stock Price Prediction Using the ARIMA Model, March 2014.
- [3] Jingyi Shen and M. Omair Shafq, Short-term stock market price trend prediction using a comprehensive deep learning system, 2020.
- [4] Charalampos M. Liapis, Aikaterini Karanikola, and Sotiris Kotsiantis, A Multi-Method Survey on the Use of Sentiment Analysis in Multivariate Financial Time Series Forecasting, 2021 Nov 29.