BRAIN. Broad Research in Artificial Intelligence and Neuroscience

e-ISSN: 2067-3957 | p-ISSN: 2068-0473

Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Center, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)

2024, Volume 15, Issue 3, pages: 153-166.

Submitted: August 5th, 2024| *Accepted for publication:* October 2nd, 2024

Deep Learning for Financial News Analysis and Stock Price Prediction: A Case Study of TCS

Bijesh Dhyani [1]
Sanjay Taneja [2]
Chandra Prakash [3]
Rajesh Tiwari [4]
Ercan Özen [5]

[1] Department of Management Studies, Graphic Era (Deemed to be University), Dehradun, Bharat (India), bijeshdhyani.mgt@geu.ac.in
[2] Department of Management Studies, Graphic Era (Deemed to be University), Dehradun, Bharat (India), drsanjaytaneja1@gmail.com, https://orcid.org/0000-0002-3632-4053
[3] Department of Management Studies, Graphic Era (Deemed to be University), Dehradun, Bharat (India), chandra.thukral.19@gmail.com
[4] Department of Management Studies, Graphic Era (Deemed to be University), Dehradun, Bharat (India), ambitioncfarajesh@gmail.com, https://orcid.org/0000-0002-5345-2508
[5] University of Uşak, Faculty of Applied Sciences, Dept. of Finance and Banking Türkiye, ercan.ozen@usak.edu.tr, https://orcid.org/0000-0002-7774-5153

Abstract: *Finance has been the most important aspect of people's lives for making small to large decisions in life. Finances affect our ability to invest in opportunities, save for uncertain future events, and purchase necessities. In addition, money has a significant impact on reducing poverty, creating jobs, and advancing society. On the other side, deep learning is a growing field as it is transforming and revolutionizing various areas and industries. While realising the importance of both finance and deep learning in our lives for better decision-making, our study aims to find the connectedness between both so that understanding this relation and utilising it will create better systems for optimal decision-making and also to study its influence on the stock market, presenting various challenges and opportunities regarding this matter and unleashing the potential that comes with the advanced technologies.*

Keywords: *deep learning; finance; stock price prediction; market news analysis; TCS*

How to cite: Dhyani, B., Taneja, S., Prakash, C., Tiwari, R., & Özen, E. (2024). Deep learning for financial news analysis and stock price prediction: A case study of TCS. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*, 15(3), 153-166. https://doi.org/10.70594/brain/15.3/11

1. Introduction

Finance being a necessity helps in deciding day-to-day transactions and activities starting from needing funds to buy necessary commodities to get a loan or make an investment to meet long-term financial goals like higher studies, buying a car, and so on. Nowadays, with the rapid upgrading of technology, technologies like deep learning have led to the transformation in the finance sector through automating various processes and fully utilising the data generated. Integrating finance and deep learning has opened up new opportunities and challenges, unprecedentedly shaping our daily lives.

As a discipline, finance encompasses the study of how people, firms, and institutions handle their assets and income. It involves personal finance, corporate funding, and government budgeting. "For individuals, finance impacts their capacity to obtain credit and loans, which in turn allows them to acquire goods and services that might otherwise be beyond their reach. Nonetheless, an overabundance of debt can result in financial strain and negatively impact one's overall quality of life (Mahapatra, Alok, & Raveendran, 2017). Furthermore, finance determines individuals' capacity to save for the future, providing a foundation for achieving long-term financial goals such as homeownership, entrepreneurship, and retirement planning. The complex dynamics of finance also extend to investment opportunities, where individuals can grow their wealth over time through strategic investment in shares, bonds, and other financial instruments. Nevertheless, investing involves risk, necessitating careful consideration of potential rewards and pitfalls (Chen & Volpe, 1998).

Apart from personal finances, the financial sector is essential for the economy, the creation of employment, and the reduction of poverty. The financial system lets investors transfer money to companies and governments so they may expand, develop, and deliver vital services to society. However, during instances like the global financial crisis of 2008, which led to several layoffs, bankruptcies, and financial distress, the connection between finance and society became apparent. Therefore, both people and society's growth must comprehend how money affects day-to-day living.

Deep learning models like LSTM have changed traditional banking ways. It is also the subset of AI which includes ANN (Artificial Neural Network). It has the capability of learning from large datasets to make predictions and lead to automation and streamlining of work. It is highly useful in the Finance sector for risk assessment. This creates a good opportunity for Finance professionals to utilize deep learning for better understanding the market factors, identify anomalies, and therefore make better decisions (Kanevsky et al., 2016).

The study evaluates the connection between Finance and Deep Learning, which majorly focuses on the impact of this relationship on day-to-day market activities. It examines the impact of money on an individual's decision-making skills, market fluctuations, how news related to any share affects its price, analysis of sentiment, merits, and demerits of deep learning in Finance. A thorough understanding of the relationship between these two enables many organizations and institutions to use it for better decision-making and other purposes as well.

2. Literature Review

There is a very significant association between share prices and any information available about these shares. Using advanced textual analysis, stock price changes and information were found to have a strong association, especially when considering the tone of the news (McQueen & Roley, 1993). Business cycles are primarily influenced by a unique shock that doesn't immediately impact productivity or monetary factors but rather represents future technological opportunities. This shock, reflected in stock prices, precedes productivity growth and explains half of business cycle fluctuations (Cole & Ohanian, 1999). Previous studies suggest the limited impact of macroeconomic news on stock prices. However, our research reveals a stronger relationship, particularly during different business cycle stages. We also find that in a robust economy, the stock market shows pessimistic results to the news of an increase in economic activities due to larger discount rate increases compared to expected cash flows (Birz & Lott Jr, 2011). The paper finds that macroeconomic news explains only one-third of stock return variance, and market reactions to major events are small, casting doubt on the complete explanatory power of change in stock price due to news. This study investigated the correlation between the price of stocks and actions on online discussion sites. While higher online discussions correlated with unusual earnings and trading volume, it failed to forecast industry-adjusted returns, suggesting market efficiency (Cutler, Poterba, & Summers, 1988). This research looks into how news about the monetary supply, the rate of inflation, real activity in the economy, and interest rates affect stock prices. The results validate the efficient markets theory by demonstrating that stock prices are only influenced by unforeseen events, especially those of monetary policy decisions. The impact of inflation surprises is found to have little evidence, real activity surprises are found to have no evidence, and stock price reactions besides the day of announcements are found to have limited support (Pearce & Roley, 1984).

The study aims to enhance computer-aided forecast precision for the stock market in Hong Kong by creating a system that incorporates stock prices and marketplace information. By using multi-kernel learning, the system achieves better directional accuracy compared to baseline systems that rely on a single information source or simple integration methods (Li et al., 2011). The paper examines how stock prices react to 15 macroeconomic announcements. Monetary variables have the greatest impact, especially on financial company stocks. The findings suggest that the market views the Federal Reserve as influential in future macroeconomic trends. Changing the operating target of the Federal Reserve after October 1982 did not significantly alter stock price responses, but it did affect short-term interest rates (Menike, 2006). IR firms tend to generate more positive press coverage, leading to higher announcement returns. However, around earnings announcements, this effect diminishes, suggesting that positive media coverage raises expectations, leading to disappointment when confronted with concrete information. The study also suggests a causal relationship between IR firms and both media coverage and stock returns, using reporter connections and geographical links as evidence (Solomon, 2012).

The output of Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) was evaluated in capturing semantic and context information from data. The findings indicated that CNN performed better in capturing semantic information from news titles, while RNN excelled in modelling complex temporal characteristics (Vargas, De Lima, & Evsukoff, 2017). The use of BERT for emotion recognition in news composition is proposed to provide relevant facts for making judgments in the stock market (O'Neill & Harcup, 2009). The strengths and weaknesses of algorithms of deep learning are analysed for stock market forecasting and examined the impact of three unsupervised variable selection approaches on the capacity of the network to predict upcoming market movements (Chong, Han, & Park, 2017). Discussed the utilization of deep learning architectures such as Multilayer Perceptron, RNN, LSTM, and CNN for forecasting share prices based on stock price data from the NSE and the NYSE. The neural network models outperformed traditional linear models like ARIMA (Yadav, Jha, & Sharan, 2020). On Comparing four forecasting models; SVM, ANN, random forest, and Naive-Bayes using two different approaches for input data representation. One approach involved quantifying ten operational variables using share trading data, while the other focused on portraying these indicators as time series data (Xiao et al., 2015). Employed deep learning models, including Paragraph Vector and Long Short-Term Memory (LSTM), for temporal data anticipation in finance (Chen, Liao, & Hsieh, 2019). Examined the connections between mass sentiment on Twitter and changes in the share prices of a company. It was found that there was a significant connection between share prices and the overall sentiment expressed in tweets (Rao & Srivastava, 2012). Enhance the accuracy of share price forecasts by analyzing a large dataset of temporal data and related press reports using deep learning models (Mohan et al., 2019). Discussed the design, implementation, and evaluation of a framework that utilized data from web forums to forecast future share prices through sentiment classification using machine learning techniques (Jelodar et al., 2020). Focused on forecasting share price movements using sentiment from web forums by incorporating specific topics related to the company, which were automatically extracted using proposed and existing topic models (Molnár et al., 2015). Utilized deep learning architectures to identify potential dynamics in data. They applied a window-sliding approach for short-term price forecasting of NSE-listed firms and the performance of three different deep learning frameworks; LSTM, RNN, and CNN were compared using percentage error as a metric (Kim & Won, 2018).

3. Data and Methodology

The dataset used in this study comprises 104,000 rows of news articles collected exclusively for the year 2022. The articles were sourced from two platforms: The New York Times's website and Google News for sentiment analysis and stock data is being collected from Yahoo Finance for further study of TCS stock by using the LSTM network. Each article for sentiment analysis in the dataset contains keywords related to TCS, banking, finance, the stock market, and the economy. The data includes the URL, heading, content, and publish date of each article. The publish date is utilized to align the news data with the corresponding financial time series. The proposed model focuses on utilizing the news titles and content as input to capture the most relevant information for analysis.

Table 1. News Data

Date	Title	Compound	Positive	Negative	Neutral	Subjectivity
01-01-2022	"15 Best Car Books for Dad For Father's Day - N..."	0.637	0.3	0	0.7	0.3
01-02-2022	"Eugene McCormack had time for everyone, says f..."	0.494	0.2	0	0.8	0
01-03-2022	"Original Content podcast: â€˜The Old Guardâ€™ ..."	0.315	0.3	0.2	0.5	0.41
01-04-2022	"Report: Jaguar Land Rover linked with UK battle..."	0	0	0	1	0
01-05-2022	"Marvel Comics solicitations January 2022: Wolv..."	0.422	0.2	0	0.8	0
27-12-2022	"Serial NYC shoplifter who's dodged jail nabbed..."	0	0	0	1	0.23
28-12-2022	"Buffalo couple cares for stranger's body after..."	0.459	0.2	0	0.8	0.25
29-12-2022	"Twitter back online after global outage hits t..."	0	0	0	1	0.15
30-12-2022	"Amazon Prime Video New Releases: December 2022..."	-0.03	0.1	0.2	0.7	0.45
31-12-2022	"NYC murders down but major crimes surge as 202..."	-0.36	0	0.2	0.8	0.41

For performing Time Series Analysis, we used TCS (Tata Consultancy Services) stock. We chose TCS stock because of the intention that we might get a large volume of news-related data easily about the company. We have used the collected data to reach the desired output and used seven technical indices which will help in building our model.

Table 2. Share Price of TCS

Date	Open	High	Low	Close	Adj Close	Volume
31/1/22	9.74	10.2	9.6	10.2	10.2	467500
1/2/22	10.1	10.2	9.9	10.2	10.18	436100
2/2/22	10.1	10.4	9.8	10	10.01	574200
3/2/22	9.81	9.89	9.6	9.63	9.63	468100
4/2/22	9.56	10	9.4	9.93	9.93	714400
7/2/22	9.98	10.2	9.9	9.94	9.94	672500
8/2/22	9.96	10.6	9.9	10.5	10.53	1117400
9/2/22	8.2	8.72	7.8	8.13	8.13	6846200
10/2/22	8.02	8.38	8	8.05	8.05	1757400
11/2/22	8.47	8.77	8.4	8.63	8.63	2065700

The stock market faces high volatility owing to numerous reasons like elections, political instability, national and international news, etc. The market may be greatly impacted by budgetary updates, which include updates on the inflation, growth, interest, and rate of employment. New legislation and political shifts may directly affect some industries or companies, which might result in a decline in share values. Adverse consequences can also result from widespread occurrences like global outbreaks and natural catastrophic events. The price of shares is also impacted by business information, including mergers and acquisitions, revenue, and other press releases. Investors must acknowledge the vital significance of these new instances.

The properties of the dataset and the relationship between its variables may be understood by using statistical measurements, analysis of factors, and analysis of correlation function. Comprehensive reports with statistical measurements, graphical representations, and correlation matrices are produced by programmes like Pandas which will lead to further accuracy in our analysis.

4. Model Design

To predict the future value of shares, we have integrated information analysis with TCS stock price data in our model. To generate a complete dataset, we first do sentiment analysis on the gathered data to ascertain the general sentiment, and then we combine it with TCS stock price data. We use a neural network framework to predict future share values. Over a long period, neural networks are effectively capturing complex relationships and patterns in the datasets. By using past data to train it, neural networks can recognise underlying trends and forecast share values in the future. To assess the neural network's predicting accuracy, the output is contrasted with the real share prices.

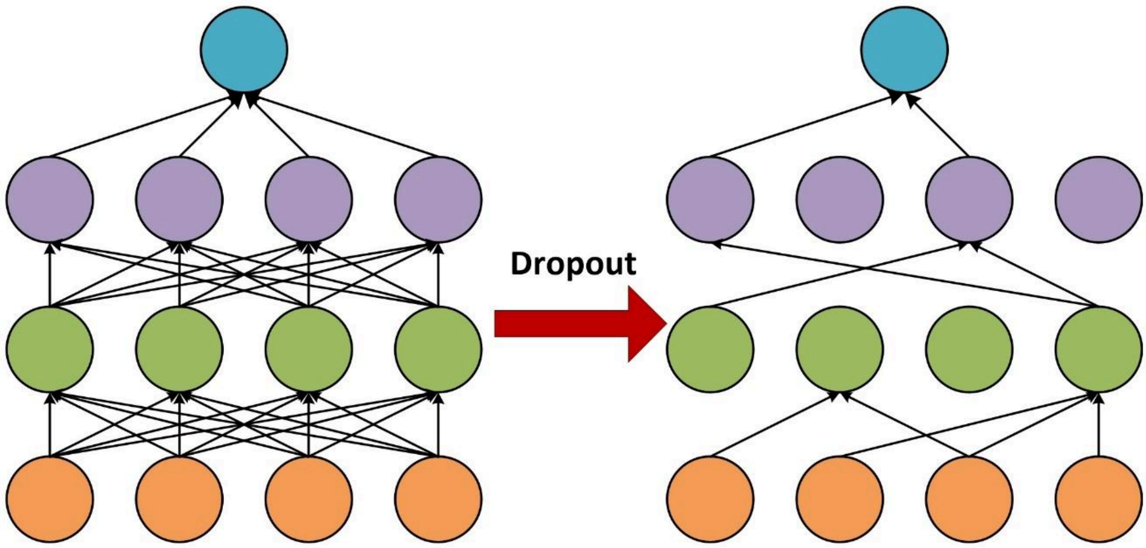


Figure 1. Neural network flowchart

By including sentiment analysis in the model as an additional layer of information, we can understand the fluctuations in the market and the feelings of investors as reflected in the new data. Knowing the expectations and basic perceptions of market participants helps explain the fluctuations in share prices. Though neural networks have a lot of potential there is always a need to also keep in mind its limitations. The quality and quantity of datasets, network structure, and other factors influence the success rate of the model.

Still, caution is necessary, and the model's projections should be backed up by a careful analysis of other factors including shifts in market trends, indicators of economic health, and firm-specific data (Zhang, 2003). In short, sentiment analysis, news analysis, and neural network programming work together to deliver relevant data for optimal decision-making about their investment in TCS stock. It's a dynamic approach that leverages information and technological advancements more effectively to improve the relevancy of predictions and our understanding of share price volatility.

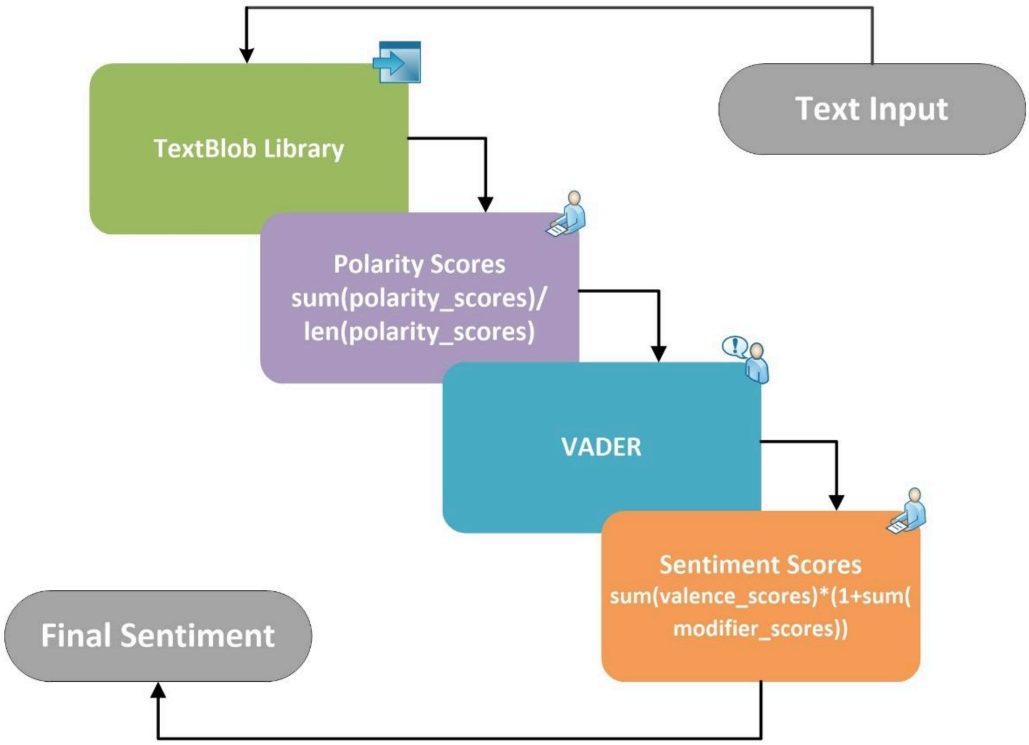


Figure 2. Process of sentiment analysis

VADER, employs a lexicon-based system, with a predetermined collection of phrases and expressions matched with sentiment ratings. Thus, it is a suitable fit for analysing brief data, like posts on Facebook and Twitter. It provides additional characteristics, such as the degree of content as either favourable or adverse, and is extremely accurate in detecting emotions. Speed is the criterion that makes it relevant for sentiment analysis on a huge scale. VADER needs help and support while dealing with humorous language and works poorly with peculiar information (Sohangir, Petty, & Wang, 2018). VADER and Text Blob are technological innovations that can provide critical insights into investor views and market trends for financial data analysis. It is very important to use both the above tools together and to be used with caution while results are comprehended. They offer distinct views on sentiment analysis. Text Blob is famous as compared to Vader because of its precision and ease of use. VADER is powerful when it comes to customisation and the ability to handle complexities in natural language. Both of them need to be considered to assess the investment opportunity for optimal decisions.

5. Share Price Forecast using Deep Learning

The complexity of data for predicting stock prices can be properly handled by deep learning and neural networks due to their robustness and capability to establish the associations and patterns between variables. The historical data, sentiment scores, and related relevant data can be processed by the neural network to predict the stock price of TCS. Text Blob and VADER are useful for the collection of sentiment ratings. The data must be processed with regard to making it ready to be used for analysis and prediction (Morid et al., 2020). Further, we developed a deep learning model based on neural network architecture; LSTM. It is suitable for analysis of time series data to extract hidden trends and associations in the datasets (Narayanan & Georgiou, 2013). The model is trained using data that enables to lessening of the differences between the algorithm's predictions and the actual value of the stock. We used readily available and accessible data to train the algorithm to reduce the discrepancy between the projected and actual values. Preferences and weights in NN are modified using gradient descent and backpropagation techniques in the training phase.

A Deep Learning model can predict both the residual new and untrained data once it has completed training. It makes use of sentiment scores, historical price data, and other relevant data to make predictions of the price of TCS shares for the selected period (Ding & Qin, 2020). Regardless of NN's ability to identify complex relations and patterns, it can be faulty as well and sometimes the historical data is insufficient to fully account for factors that continue to impact stock prices, such as uncertainty of the future, volatility of the market or maybe changes in the perceptions of investors (Prakash, Dhyani, & Lalit, 2024).

Forecasts for next year are obtained by using deep learning, ANN along with sentiment analysis, historical price data, and various other relevant variables. To make informed investment decisions, the results must be carefully examined and considered in along with information from other sources.

Long Short-Term Memory (LSTM)

LSTM is a neural network that is used for sequence modelling and time-series prediction tasks. It is a type of neural network that aids in getting over the problems of conventional neural networks in terms of addressing long-term dependency in sequences. By employing a memory block for long-term storage of data, LSTM addresses this problem (Nelson, Pereira, & De Oliveira, 2017). The flow of information inwards and outwards from the LSTM cell is regulated by output, forget, and input gates. The sequence context is stored in the LSTM cell. The update of the hidden state is done by using the state of the preceding cell and the previous hidden state (Landi et al., 2021). Specifying the loss function in a particular architecture along with optimisation of a model by employing a backpropagation through time (BPTT) algorithm is used to train the LSTM model. The sequences of input are fed into the model, loss is calculated followed by adjustments in the model parameters for loss minimization. The performance of the model is enhanced by the adoption of techniques like attention mechanism, batch normalisation, and drop out. The input sequence emphasis of the model is enhanced by the attention mechanism. Training speed is improved by batch normalisation. Overfitting is prevented by dropout. LSTM accommodates input sequences of variable length and manages long-term dependencies. LSTM has diverse applications. The performance of LSTM can be enhanced further by employing additional techniques. The input from the previous hidden state, the current input, and a bias term is taken by the sigmoid function (Landi et al., 2021). A schematic diagram of LSTM is shown in Figure 6.

The information from the previous hidden state that needs to be forgotten is selected by the forget gate, as shown in the equation below

$$f_t = \sigma(W_f * [h_t - 1, x_t] + b_f) \quad (6)$$

The weight matrix is represented by W_f, the previous hidden state is shown by $h_t - 1$, the current input is represented by x_t, bias term is represented by b_f.

The input gate determines fresh information to be presented to the memory cell. The input gate is shown by the below-mentioned equation.

$$i_t = \sigma(W_i * [h_t - 1, x_t] + b_i) \quad (7)$$

bias term is represented by b_i and weight matrix is represented by W_i

The output gate determines the proportion of the memory cell to the next hidden level and to be output which is shown by the equation below.

$$o_t = \sigma(W_o * [h_t - 1, x_t] + b_o) \quad (8)$$

bias term is represented by b_o and weight matrix is represented by W_o

The combination of the input gate and forget gate updates the memory cell and is given by the below-mentioned equation.

$$C_t = f_t * C_t - 1 + i_t * tanh(W_c * [h_t - 1, x_t] + b_c) \quad (9)$$

the weight matrix is represented by W_c, the previous memory cell is represented by $C_t - 1$

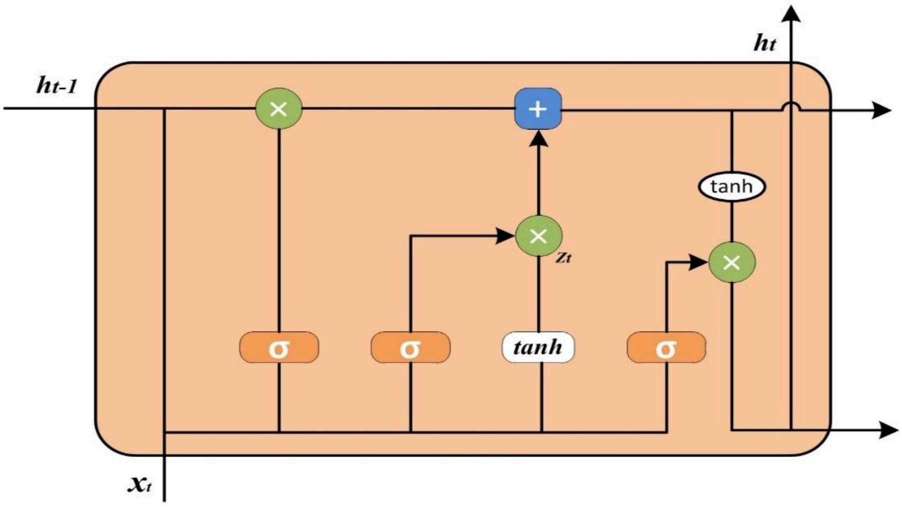
The updated memory cell passes through the output gate and function of activation to arrive at a hidden state. The hidden state is a hyperbolic tangent function:

$$h_t = o_t * tanh(C_t) \quad (10)$$

The output gate is represented by o_t, the updated memory cell is represented by C_t, and the hyperbolic tangent function is represented by tanh.

The LSTM algorithm for capturing dependencies of long-term nature in sequential data has been demonstrated (Siami-Namini, Tavakoli, & Namin, 2018).

Figure 6. LSTM schematic diagram



6. Model Evaluation

The accuracy of the model is exhibited by reliability, interpretations, and generalisations. This demonstrates the decision-making of the model and reliability tests the resistance of the model towards disruptions and damages. A combination of qualitative incorporating expert evaluation, user research, and multiple factor validation along with quantitative techniques ensures robust evaluation of the model (Narayanan & Georgiou, 2013). Various measures of error like squared mean, root of squared mean, and mean of absolute error are used for evaluation of performance (Prakash, Barthwal, & Acharya, 2022).

6.1. Mean Squared Error (MSE)

The average of the squared difference between predicted values and actual is depicted by MSE. The measure can detect the outliers (Bamana, Kamalabad, & Oberski, 2024).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$$

Observations are represented by n, and actual and predicted values are represented by y_i and $\hat{y}_i$ respectively.

6.2. Root Mean Squared Error (RMSE)

It is a useful tool as it is easier to interpret. This tool is useful as it is also able to predict outliers (Iftikhar et al., 2024).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$$

6.3. Mean Absolute Percentage Error (MAPE)

This tool is useful to understand the prediction accuracy in percentage, making it useful for analysis. However, this tool has the limitation of processing data when actual values are near zero, creating large percentage errors (Sohrabi, Jodeiri Shokri, & Dehghani, 2023).

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\%$$

These tools provide the results in different aspects of units of measurement, MAPE provides relative comparison while MSE gives results as square units, and RMSE provides results in the same result (Albeladi, Zafar, & Mueen, 2023). The fitness of the regression model is demonstrated by R squared which depicts the degree of variance explained by the independent variable.

Root Mean Squared Error (RMSE), which is frequently used to assess and compare the outcomes of various models, is the square root of the Mean Squared Error, the difference between the expected and actual values.

Though these metrics provide useful insights, they should be evaluated in light of the specific scenario and used in conjunction with other features. Understanding their constraints is crucial, as is taking into account the model's intended purpose and data features. A thorough assessment procedure allows us to ensure that models are exact, trustworthy, and suitable for the purposes for which they have been built.

7. Results and Discussion

In this study, we introduced a model that combined the news related to TCS share price from Yahoo Finance with news collected from Google News, and the NY Times. Using the widely used libraries VADER library and Text Blob, we conducted sentiment analysis and looked at the relationship between changes in share prices and news opinion. Based on the sentiment data from the news, a neural network framework was selected to predict the highest and lowest share values of TCS.

The neural network model's dependability was assessed using the mean squared error. The obtained mean squared difference- MSE between the actual and predicted share prices was 2.49. The square root of the mean square error or 1.57, is the root mean square error, which quantifies the average size of forecasting errors. The average percentage difference between the actual and predicted share prices was calculated using the Mean Absolute Percentage Error, which produced a result of 0.22.

Moreover, 0.0066 was discovered to be the R-squared value, which establishes the degree of variation explained by the model. This implies that the news sentiment data can only account for a small portion of the fluctuation in share prices.

Being mostly based on news sentiment data, the performance evaluation metrics show that the neural network framework is not able to forecast TCS share prices, whether high or low. There is a difference between the actual price and the predicted price, as observed in the low value of variance. The results must be used with care and by considering macroeconomic variables and market information that will have an impact on stock prices. Further research is needed to enhance the robustness of the model.

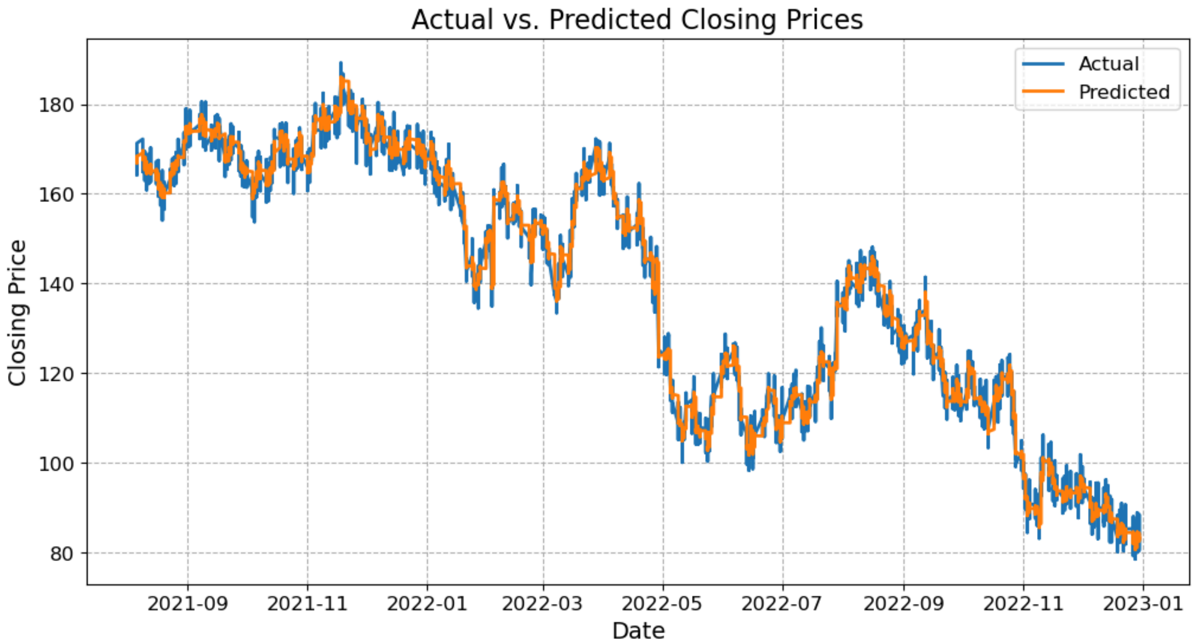


Figure 3. Forecast graph of the neural network

Table 3. Model Summary

Model	MSE	MAPE	RMSE
Neural network	2.49	0.226	1.578

Model Validation

Reliability and accuracy are assured by model validation (Veldkamp et al., 2018). Though traditionally estimation of error values is done to assess the validity of a model, it has limitations when the data and model are complex. The ability to factor in new and unseen data is a challenge for modelling, more so in the case of prediction of the price of a stock. To overcome the challenge, it is attempted to develop the model on each small subset of data to achieve cross-validation. Table 4 shows the model validation using a comparison with existing work.

Table 4. Model validation using comparison with existing work

Study	Data source	Classification technique	Attributes	Model performance
(Chen, Liao, & Hsieh, 2019)	DataHub	Sentiment Analysis	Title, Content and Publishing date, Stock data	–
(Rao & Srivastava, 2012)	Kaggle	ARIMA	Stock data	–
(Mohan et al., 2019)	MoneyControl	DNN	Deep learning automatically extracts features from data	MSE-0.9655 MAE-0.5806
(Jelodar et al., 2020)	NSE	ANN	Stock data	MAPE-4.88

(Molnár et al., 2015)	NasDAQ	ANN, SVM, random forest, and Naive-Bayes	Stock data	MAPE - 5.00 MAE - 2.06 RMSE -108.64 MSE- 19013.75
(Kim & Won, 2018)	EODData	Sentiment Analysis	Compound of sentiment, Stock data	–
(Zhang, 2003)	Twitter data	Sentiment Analysis	Twitter Sentiments, Stock data	MSE-0.632
Our approach	Yahoo Finance	Sentiment Analysis, and LSTM	New York Times's website, Google News, Yahoo Finance	MSE-2.49 MAPE-0.226 RMSE-1.578

8. Conclusion

Aiming to unleash the potential of today's time technology, our paper was centered on understanding the connection between share price movements of TCS' shares and the emotions of people with news. Even though our model used neural networks and sentiment analysis for forecasting, the assessment findings showed constraints in reliably predicting share price values based on data on news sentiment alone. R2 is used for performance evaluation which showed that a lower value of R2 further means that the model explained only a small part of deviations in the TCS share price. Thus, it is advisable not to blindly follow the model's results for making decisions before investing rather investors should consider various other factors that impact the share price movements like economic and market conditions, industry information, and so on for better decision-making. To improve the model's effectiveness and broaden our knowledge of the connection between investors' reactions to news and share prices of TCS, more investigation and analysis are required. In general, a thorough strategy that incorporates a variety of techniques and variables is required to make informed investing decisions when taking TCS stock price estimates into account.

References

Albeladi, K., Zafar, B., & Mueen, A. (2023). Time Series Forecasting using LSTM and ARIMA. International Journal of Advanced Computer Science and Applications, 14(1), 313-320.

Bamana, A. B., Kamalabad, M. S., & Oberski, D. L. (2024). A systematic literature review of time series methods applied to epidemic prediction. Informatics in Medicine Unlocked, 101571.

Birz, G., & Lott Jr, J. R. (2011). The effect of macroeconomic news on stock returns: New evidence from newspaper coverage. *Journal of Banking & Finance*, *35*(11), 2791-2800.

Chen, H., & Volpe, R. P. (1998). An analysis of personal financial literacy among college students. *Financial services review*, *7*(2), 107-128.

Chen, M. Y., Liao, C. H., & Hsieh, R. P. (2019). Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach. *Computers in Human Behavior*, *101*, 402-408.

Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications*, *83*, 187-205.

Cole, H. L., & Ohanian, L. E. (1999). The Great Depression in the United States from a neoclassical perspective. *Federal Reserve Bank of Minneapolis Quarterly Review*, *23*, 2-24.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1988). *What moves stock prices?* (Vol. 487). Cambridge, Massachusetts: National Bureau of Economic Research.

Ding, G., & Qin, L. (2020). Study on the prediction of stock price based on the associated network model of LSTM. International Journal of Machine Learning and Cybernetics, 11(6), 1307-1317.

Iftikhar, H., Gonzales, S. M., Zywiołek, J., & López-Gonzales, J. L. (2024). Electricity demand forecasting using a novel time series ensemble technique. IEEE Access.

Jelodar, H., Wang, Y., Orji, R., & Huang, S. (2020). Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach. *IEEE Journal of Biomedical and Health Informatics*, *24*(10), 2733-2742.

Kanevsky, J., Corban, J., Gaster, R., Kanevsky, A., Lin, S., & Gilardino, M. (2016). Big data and machine learning in plastic surgery: a new frontier in surgical innovation. *Plastic and reconstructive surgery*, *137*(5), 890e-897e.

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, *103*, 25-37.

Landi, F., Baraldi, L., Cornia, M., & Cucchiara, R. (2021). Working memory connections for LSTM. Neural Networks, 144, 334-341.

Li, X., Wang, C., Dong, J., Wang, F., Deng, X., & Zhu, S. (2011). Improving stock market prediction by integrating both market news and stock prices. In *Database and Expert Systems Applications: 22nd International Conference, DEXA 2011, Toulouse, France, August 29-September 2, 2011, Proceedings, Part II 22* (pp. 279-293). Springer Berlin Heidelberg.

Mahapatra, M. S., Alok, S., & Raveendran, J. (2017). Financial literacy of Indian youth: A study on the Twin Cities of Hyderabad–Secunderabad. *IIM Kozhikode Society & Management Review*, *6*(2), 132–147. https://doi.org/10.1177/2277975216667096

McQueen, G., & Roley, V. V. (1993). Stock prices, news, and business conditions. *The review of financial studies*, *6*(3), 683-707.

Menike, L. M. C. S. (2006). The effect of macroeconomic variables on stock prices in emerging Sri Lankan stock market. *Sabaragamuwa university journal*, *6*(1), 50-67.

Mohan, S., Mullapudi, S., Sammeta, S., Vijayvergia, P., & Anastasiu, D. C. (2019, April). Stock price prediction using news sentiment analysis. In *2019 IEEE fifth international conference on big data computing service and applications (BigDataService)* (pp. 205-208). IEEE.

Molnár, S., Molnár, M., Naár-Tóth, Z., & Timár, T. (2015). Forecasting share price movements using news sentiment analysis in a multinational environment. *HUNGARIAN AGRICULTURAL ENGINEERING*, (28), 53-55.

Morid, M. A., Sheng, O. R. L., Kawamoto, K., & Abdelrahman, S. (2020). Learning hidden patterns from patient multivariate time series data using convolutional neural networks: A case study of healthcare cost prediction. *Journal of Biomedical Informatics*, *111*, 103565.

Narayanan, S., & Georgiou, P. G. (2013). Behavioral signal processing: Deriving human behavioral informatics from speech and language. *Proceedings of the IEEE*, *101*(5), 1203-1233.

Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In 2017 International joint conference on neural networks (IJCNN) (pp. 1419-1426). Ieee.

O'neill, D., & Harcup, T. (2009). News values and selectivity. In *The handbook of journalism studies* (pp. 181-194). Routledge.

Pearce, D. K., & Roley, V. V. (1984). *Stock prices and economic news* (No. w1296). National bureau of economic research.

Prakash, C., Barthwal, A., & Acharya, D. (2022). FLOODWALL: a real-time flash flood monitoring and forecasting system using IoT. IEEE Sensors Journal, 23(1), 787-799.

Prakash, C., Dhyani, B., & Lalit, N. (2024). A Comprehensive Study on Structural and Non-Structural Databases and Its Impact on Hybrid Databases. In Navigating the World of Deepfake Technology (pp. 351-377). IGI Global.

Rao, T., & Srivastava, S. (2012). Analyzing stock market movements using twitter sentiment analysis.

Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018, December). A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE international conference on machine learning and applications (ICMLA) (pp. 1394-1401). Ieee.

Sohangir, S., Petty, N., & Wang, D. (2018, January). Financial sentiment lexicon analysis. In *2018 IEEE 12th international conference on semantic computing (ICSC)* (pp. 286-289). IEEE.

Sohrabi, P., Jodeiri Shokri, B., & Dehghani, H. (2023). Predicting coal price using time series methods and combination of radial basis function (RBF) neural network with time series. Mineral Economics, 36(2), 207-216.

Solomon, D. H. (2012). Selective publicity and stock prices. *The Journal of Finance*, *67*(2), 599-638.

Vargas, M. R., De Lima, B. S., & Evsukoff, A. G. (2017, June). Deep learning for stock market prediction from financial news articles. In *2017 IEEE international conference on computational intelligence and virtual environments for measurement systems and applications (CIVEMSA)* (pp. 60-65). IEEE.

Xiao, L., Wang, J., Dong, Y., & Wu, J. (2015). Combined forecasting models for wind energy forecasting: A case study in China. *Renewable and Sustainable Energy Reviews*, *44*, 271-288.

Yadav, A., Jha, C. K., & Sharan, A. (2020). Optimizing LSTM for time series prediction in Indian stock market. *Procedia Computer Science*, *167*, 2091-2100.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175.