Predict financial text sentiment: an empirical examination

Ruchi Kejriwal, Monika Garg and Gaurav Sarin
Delhi School of Business, Pitampura, New Delhi, India

Received 26 June 2022
Revised 16 September 2022
Accepted 31 October 2022

Abstract

Purpose – Stock market has always been lucrative for various investors. But, because of its speculative nature, it is difficult to predict the price movement. Investors have been using both fundamental and technical analysis to predict the prices. Fundamental analysis helps to study structured data of the company. Technical analysis helps to study price trends, and with the increasing and easy availability of unstructured data have made it important to study the market sentiment. Market sentiment has a major impact on the prices in short run. Hence, the purpose is to understand the market sentiment timely and effectively.

Design/methodology/approach – The research includes text mining and then creating various models for classification. The accuracy of these models is checked using confusion matrix.

Findings – Out of the six machine learning techniques used to create the classification model, kernel support vector machine gave the highest accuracy of 68%. This model can be now used to analyse the tweets, news and various other unstructured data to predict the price movement.

Originality/value – This study will help investors classify a news or a tweet into "positive", "negative" or "neutral" quickly and determine the stock price trends.

Keywords Financial markets, Text mining, Sentiment analysis, Classification techniques, Predictive analysis

Paper type Research paper

Introduction

Stock market is a marketplace or a platform where shares of various public listed companies are available and buying and selling. Prices of shares can be determined by various factors; one of the main criterion is demand and supply. There are various stock market exchanges and are controlled by some regulatory body. In India, the Bombay Stock Exchange (BSE) and National Stock Exchange (NSE) are regulated by The Securities and Exchange Board of India (SEBI). Similarly, in the USA, New York Stock Exchange (NYSE) and the Nasdaq are regulated by Securities and Exchange Commission (SEC). Investors invest in stock markets for higher returns, but there are various factors that affect the market, making it volatile and thus risky. Other than supply and demand, various factors that affect the stock market are company-related factors, interest rates, politics, exchange rates, sentiments etc. Market involves risks, although the risk cannot be eliminated, but it can be minimised. The risk can be controlled by understanding the market and making the requisite analysis. Investors generally analyse the market to reduce the risk, which can be either done through fundamental analysis or technical analysis. Market is analysed by investors based on

Vilakshan – XIMB Journal of Management
Emerald Publishing Limited
0973-1954
DOI 10.1108/XJM-06-2022-0148

fundamentals, which indicates the business performance or technical indicators that are usually based on feeling and emotions. Technical indicators are used to analyse the trends and short-term price movements to earn profits. These trends are often impacted by market sentiments, making it necessary to understand the sentiments to earn better profits. Market sentiments is the collective attitude or psychology of the investors toward the financial market that determine the price of the assets. Market sentiment is either bullish or bearish. The market is said to be bullish when the economy is performing better and the share prices are rising, and the market is said to be bearish when the economy is impacted due to several reasons and share prices start falling. Also, due to rapidly increasing unstructured data along with importance of market sentiments, it becomes necessary to analyse the financial sentiments to take informed and better decisions. Unstructured data is collected through e-mails, blog posts, social media channels, comments and so on.

Analysing the sentiments means to determine the emotional tone or determining the attitude of the text and classify it as "positive", "negative" or "neutral". Doing this task manually for the numerous data available can be a tedious task; hence, with the help of data mining, machine learning and artificial intelligence, the emotion of the text is determined, which is referred to as sentiment analysis. Sentiment analysis is the process of using natural language processing (NLP). Sentiment analysis is a powerful tool to analyse the unstructured data. This is used for managing the brand reputation by understanding the customer reviews and perception, for market analysis to ensure better decision-making, used in politics, evaluating the success of the marketing campaign, analysing the financial sentiments. For this purpose, there are various types of sentiment analysis, which are categorised as follows:

- Fine-grained sentiment analysis: In this, the text is categorised as "positive", "negative" and "neutral".
- Emotion detection: This helps to detect various emotions like happiness, anger, shock, frustration, etc.
- Intent-based analysis: It is used to understand an intent behind a text to take corrective action. Say, a comment about faculty product delivery will require the company to either exchange the product or provide service.
- Aspect-based analysis: This helps to understand the specific component of the text mentioned. Suppose a customer complaint about poor sound quality of a television, then only the sound system needs attention, and the review for the whole product is not negative.

Understanding humans is a complex process, and understanding human sentiments can involve a few challenges. Challenges of using sentiment analysis can be as: a text can give contradictory sentiments and can be both negative and positive, a name in a text can have different meaning and is difficult to analyse the context, it may get difficult to determine the context of the pronoun, it cannot identify sarcasm, it may get difficult to analyse the abbreviations and acronyms.

Sentiment analysis is done to analyse the situation of the financial market, which is referred to as financial sentiment analysis. Predicting the sentiments of various unstructured data and the financial news will automate the task of understanding the sentiments of the market. This will in turn help the investors take quick and better decisions in the stock market to earn more profits and reduce the risk. Hence, through this research, various classification models will be created, and their accuracy will be checked, which will

thus help the investors of the stock market to take an informed and better decision by predicting the stock prices using sentiment analysis.

Motivation

With the exponential growth of unstructured data has derived the need to understand and analyse it. Financial sentiments have been observed to affect the market. This creates a need to understand the sentiments, thus enabling to understand the market better. Sentiment analysis is a powerful tool that enables investors to understand emotions. It is also useful in monitoring unstructured data, which is collected through e-mails, blog posts, social media channels, and allows to gain wider opinion on how people think. Shifts in sentiment analysis have been shown to correlate with the change in stock market. Understanding the psychology of sentiments means to determine the emotional tone or determining the attitude of the text and classify it as "positive", "negative" or "neutral" will help investors to take informed and precise data-driven business decisions.

Research objectives

To understand sentiments through unstructured data.

- To formulate various financial sentiments predictive models using classification methods.
- To check the accuracy of the models and identify the most accurate model.

Research questions

RQ1. Does unstructured data such as news and social media have an impact on the stock market?

RQ2. Is there any impact of investor's sentiments and emotions on price fluctuation in the stock market?

RQ3. How can the sentiments be studied and with how much accuracy?

Contribution

The research offers various classification models that can be used to study the sentiments behind the news or tweets. These models can be used to study the unstructured data and make informed business decisions in the stock market.

Literature review

Bhardwaj *et al.* (2015) have defined the importance of sentiment analysis in today's business and stock market. They have used sentiment analysis to predict the future prices of Indian Stock Market indices, Sensex and Nifty. Analysis is done through various machine learning and lexicon-based approaches. They have created the models that can be further used to predict the stock market prices. Li *et al.* (2016) research is done to check the effectiveness and ability of the tensor-based information framework because it can capture the multiple information on stock movements and can also improve underperforming financial products. This research also seeks to check whether social sentiments and new articles affects or change their views or attitudes towards investments. And, it has been found out that social media influences stock movement changes and has a great influence. Picasso *et al.* (2019) have done their work to combine technical and fundamental analysis approaches to predict

the market with the use of machine learning techniques and sentiment analysis. Later, they also aim to develop a robust model that helps to predict the trends of stock. Thus, the outcome of this study is that a robust model can effectively classify both positive and negative trends in portfolio and under study stocks.

Sohangir *et al.* (2018) in their paper seek to determine if the sentiment of the mass group sentiments affects the stock prices. Using deep learning methods, several model of long short-term memory (LSTM), doc2vec and convolutional neutral networks get a better understanding of sentiment analysis and whether the models are effective or not. The results concluded that convolutional neural networks performed better than other models. Similarly, Sagala *et al.* (2020) in their paper conducted various experiments and research to evaluate stock price movement classification, consisting of technical analysis feature and online medial sentiment analysis. And, it is has been concluded that the highest accuracy is achieved when a combination of features is applied and proves that sentiment analysis can also generate profits for investors because sentiment analysis can also help us to make better decision. A similar hybrid model was formulated by Jing *et al.* (2021) including technical indicators and sentiment analysis for predicting the closing price of stock for the next day. They found that this new hybrid model gave better results that the other model.

Yadav *et al.* (2020) seek to determine the sentiment strength of the news to ensure effective decision-making. For this, the authors have taken into consideration three techniques, namely, Turney's approach, hybrid approach and noun–verb approach. It was concluded that the hybrid and noun–verb approach gave better results than Turney's approach. Earlier, *Joshi* et al. *(2016)* worked on the assumption that news articles have an impact on the stock market and created three classification models: RF (random forest), support vector machine (SVM) and naïve. It was observed that RF, SVM and naïve Bayes gave an accuracy of 92, 86 and 83%, respectively. Also, Shuhidan *et al.* (2018) studied the sentiment analysis of financial new headlines in Malaysia by using machine learning algorithm techniques, i.e. opinion lexicon-based algorithm and naïve Bayes. Jangid *et al.* (2018) in their paper seek to determine the framework of deep learning models to do aspect-based sentiment analysis on financial news and headlines. In this, deep learning models are classified into three discrete classes, which are positive, negative and neutral. And, it has been found that they are able to design the model that consists of two neutral networks, one for the aspect extraction and the other for sentimental score prediction. In a similar study, Souma *et al.* (2019) have used the artificial intelligence and deep learning to enhance financial market casting results. This means we can predict the change in stock price by just observing and analysing the sentiments through quantitative data. Hence, this it clarifies that investors can interpret the polarity of news and can predict the change in price before taking and action and making investment.

Valle-Cruz *et al.* (2022) formulated a lexicon-based approach, including technical financial analysis, combined to find out the influence of twitter accounts with the behaviour of important financial indices. And, they found out that during the Covid-19 pandemic, Twitter posts have influenced financial indices more significantly than any other pandemic. Mehtab and Sen (2019) in their research have used machine learning to create to create various predictive model using classification techniques to predict the price of Nifty 50 Index based on the closing prices from 2015–2017. Further, they created an LSTM-based deep learning network to predict the stock prices.

Also, a predictive model was also created for understating the public sentiments by using data from Twitter and was compared with the market sentiment. A comparison was made between all the models, and the results show that the model using the sentiment analysis gave the best results. A similar study was done by Jin *et al.* (2020) wherein they had

used deep learning and LSTM model for stock market predictions that considers investor's emotional tendency. And the LSTM model makes stock pricing prediction simpler and easier because this model helps in focusing on the most important information of the current task. This paper's results are that the proposed scheme outperforms in three main aspects, which is closer predicted closing price, higher rise and fall classification accuracy and lower time set off. Kinyua *et al.* (2021) studied the impact of President Trump's tweets on the US stock markets indices. They used sentiment analysis, machine learning, classification and regression to analyse the market reaction on the tweets. The tweets made by the President during the open market showcased a negative reaction. Das *et al.* (2018) in their work show how sentiment analysis of public mood affected the Twitter feeds, which is used to predict the change in price of stocks in the market. Later, they also used the active learning approach, which helps them to find out more online stock data that is not get easily accessible and helps us to predict future stock prices more accurately. Thus, they can analyse the sentiments and the current stock trend in the rise or fall of prices, which will improve their reliability and trustworthiness of the future.

To accurately anticipate stock values, Gite *et al.* (2021) suggested combining machine learning methods with a deep learning method called LSTM. In Mehta *et al.* (2021), the authors demonstrated the relationship between changes in a firm's stock price and publicly voiced attitudes (opinions) about that company. They developed and applied a tool for stock price forecast accuracy that considered public opinion in addition to other factors. To predict future stock values, the suggested algorithm considered the sentiment of the general population, opinions, news and past stock prices. Machine learning and deep learning techniques, such as SVM, linear regression, naive Bayes and LSTM, were used in their research. The issue of predicting the movement of stocks and stock price indices for Indian stock markets was addressed by Patel *et al.* (2015a, 2015b). The paper compares two ways for input to these models with four prediction models:

(1) artificial neural network (ANN);
(2) SVM;
(3) RF; and
(4) naive Bayes.

The second strategy focuses on encoding these technical factors as trend deterministic data. The first approach for input data comprises computing ten technical parameters using stock trading data (open, high, low and close prices). Each prediction model's accuracy was assessed for each of the two input approaches. Reliance Industries and Infosys Ltd were evaluated using historical data spanning ten years, from 2003 to 2012, together with the CNX Nifty and S&P BSE Sensex stock price indices. The issue of forecasting future stock market index values is the main topic of the paper Patel *et al.* (2015a, 2015b). For experimental evaluation, two Indian stock market indices – CNX Nifty and S&P BSE Sensex – were used. On ten years' worth of historical data for these two indices, experiments are based. For periods of time between 1 and 10, 15 and 30, forecasts are produced. Support vector regression (SVR) is included in the first stage of the two-stage fusion approach that is proposed in the study. In the second stage of the fusion approach, SVR-ANN, SVR-RF and SVR-SVR fusion prediction models are created using ANN, RF and SVR. When compared to single-stage scenarios in which ANN, RF and SVR are applied independently, the prediction performance of these hybrid models performs better. Each of the prediction models receives ten technical indicators as inputs.

Research methodology

For financial sentiment analysis, the data set is taken from Kaggle. The data contains 5,322 unique sentences and their sentiment. The sentiment is categorised as either "positive", "negative" or "neutral". The examination of this secondary data set is done using Python. Primarily, *text mining* is done on the complete data set. Text mining, also known as text analytics, is an artificial intelligence tool that converts unstructured data (text) into structured data with the help of NLP. This is done as structured data is suitable for analysis. NLP is allowing computers to read and write by analysing sentence structure and grammar by using various methods such as computer science, artificial intelligence, linguistics and data science to understand the human language. NLP is an artificial intelligence technique that makes usage of computers more convenient.

Text mining was followed by *multiple classification technique*, which is a process of categorising the data set into two or more than two classes. The data was categorised into these classes as "positive", "negative" and "neutral". There are various methods of applying multi-class classification such as K-nearest neighbours (KNNs), decision trees, naïve Bayes, RF and gradient boosting. For this research, naïve Bayes, SVM, RF, KNN, decision tree and Kernel SVM methods are used.

The accuracy of the model created using multiple classification technique is checked using a *confusion matrix*. It is used to describe the performance of the predictive classification model. Accuracy is determined by adding all the true positives and true Negatives and then dividing it by the total test set to get the accuracy percentage. The model will the highest accuracy score will be the best out of the models created.

Data analysis

The data set consists of various statements and their sentiments as "positive", "neutral" or "negative". The data is analysed to formulate classification models to determine the sentiment of a certain news or a statement, thus enabling to understand the market. This is done through text mining, followed by creating various models using classification techniques and checking their accuracy using confusion matrix. The data was analysed using Google Colab environment. Six different classification techniques were used to make the predictive models. which are as follows:

(1) Kernel SVM;
(2) SVM;
(3) RF;
(4) Decision Tree;
(5) KNN; and
(6) Naïve Bayes.

Kernel support vector machine

This algorithm takes the input data and then transforms into the required form. It provides shortcut to avoid complications. In this, we can also go to infinite number of dimensions without raising the complexity. In this, we do not have to transform the data manually. This is done by kernel tricks. There are few kernel tricks that are linear, poly, RBF, custom. This model gave an accuracy of 68.43%. Table 1 shows the confusion matrix for the kernel SVM model.

Support vector machine

SVM is one of the best learning algorithms that is used for classification and regression problems both but majorly used for solving classification problems in machine learning. The basic concept of SVM is to create decision lines to differentiate into classes to put new data point into correct category in the future. This decision line is called hyperplane. And, the vectors that help in creating the hyperplanes are called support vector; because of that, this algorithm is named as SVM. SVM target is to find the hyperplanes with the highest margins between support vector and hyperplane.

This model gave an accuracy of 63.72%. Table 2 shows the confusion matrix for the SVM model.

Random forest

This classifier contains number of decision trees on the various parts of the given data set and then takes average of all the decision parts to improve the prediction accuracy of the data set. In RF, the prediction comes from each tree and is based on the highest number of votes of predictions, and it predicts the final output. This classifier is used because it takes less time as compared with others. RF can predict the accuracy even with the large data set or even when the large portion of data set is missing. This model gave an accuracy of 63.21%. Table 3 shows the confusion matrix for the RF model.

Decision tree

It is a supervised learning method used for both classification and regression tasks. A decision tree uses the tree representation to solve the problem in which the internal nodes represent the features of a data set, branches represent the decision rules and each leaf node represents the outcome. This classifier is easy to understand and requires little data

18	140	34
27	592	24
5	139	190
	0.6843455945252352	

Table 1. Confusion matrix – kernel SVM

43	118	31
93	473	77
19	86	229
	0.6372968349016254	

Table 2. Confusion matrix – SVM

35	122	35
88	499	56
17	112	205
	0.632164242942686	

Table 3. Confusion matrix – RF

preparation. This model gave an accuracy of 60.3%. Table 4 shows the confusion matrix for the decision tree model.

K-nearest neighbour

KNN is an algorithm that is used to solve the classification model problems. KNN tries to predict the future for the test data by comparing all the similarities between test data with the training data. Then the number of points which that are close to the test data is marked as K. Then, KNN finds the k nearest data points in the training set. This classifier is easy to implement and easily gets evolves with the new data. The model gave the accuracy of 56.2%. Table 5 shows the confusion matrix for the KNN model.

Naïve Bayes

This classifier is called Naïve because it assumes that some features are independent of any occurrence of other features, and it is called Bayes because it is dependent on the Bayes' theorem. It is one of the simplest algorithms to perform and used for solving classification problems. It is mainly used in text classification that includes large portion of the training data set. It is easy and fast machine learning, which can make quick predictions and need less training data when compared with others. This model gave an accuracy of 42.68%. Table 6 shows the confusion matrix for the naïve Bayes model.

Evaluation

The classification models were tested using the confusion matrix to determine the most effective model. Table 7 gives the accuracy of the models created using different classification techniques. The model created using kernel SVM gave the highest accuracy of 68.43%, followed by SVM and RF, with accuracy of 63.72 and 63.21%, respectively. Decision tree had an accuracy of 60.3%, followed by KNN with 56.2%. Naïve Bayes gave the

Table 4. Confusion matrix – decision tree

39	116	37
112	436	95
32	72	230
	0.6030795551753636	

Table 5. Confusion matrix – KNN

55	113	24
100	498	45
58	172	104
	0.5620188195038495	

Table 6. Confusion matrix – naïve Bayes

133	27	32
265	220	158
164	24	146
	0.4268605645851155	

lowest accuracy of 42%. Hence, the kernel SVM model can be used to analyse the statements and classify the sentiment as positive, negative or neutral.

Managerial implications

Stock price forecasting is a well-known problem. According to the efficient market hypothesis, stock prices cannot be predicted, and they act in a random walk fashion. Technical analysts, however, contend that recent stock prices reflect the majority of relevant information, making it possible to predict prices simply by observing trends in price movements. Due to the inherent uncertainty, predicting stocks and stock price indices is challenging. Before buying a stock, investors conduct two different sorts of analysis. The fundamental analysis comes first. Investors use this to determine whether or not to invest by looking at the intrinsic worth of stocks, the performance of the sector and economy, the political environment, etc. The appraisal of equities using technical analysis, on the other hand, involves looking at market data such as historical prices and volume. Technical analysts analyse stock charts to spot patterns and trends that can hint at how a stock will behave in the future rather than attempting to calculate a security's fundamental value. Investor perception of a certain company's stock at a particular period is significantly influenced by financial news. Beginning investors can learn how and when stock prices rise or fall and how to base a decision on the same by making forecasts based on news. As stock movements depend on a variety of factors, predicting future patterns in the market is a challenging undertaking. Researchers have predicted that news stories and stock prices are related, and that stock price movements may be correlated with news. This research can be used to help investors take informed decisions in the stock market. Classification models are created and are used to study the unstructured data in a quick and easy manner. Technical analysis helps investors predict the stock price movement. News and other such information have been known to impact the future stock price. Hence, this model will consider news articles and classify them into positive, negative and neutral. If the news turns out to be positive, then it will have a positive and bullish impact on the stock price movement, whereas if the news is negative, the stock can be expected to have a bearish impact. And, if the news is neutral, then there will be chances that there is no impact on the stock prices. The model automates the process of analysing the news and market sentiments and enables quick and effective data-driven decision-making.

Future scope

The maximum accuracy achieved by the model is 68%, a better model can be formulated with more accuracy. Also, the model can be tested by doing a study on various stock prices and information relating to it. This will help to get a better understanding and accuracy of the model. The present model cannot identify sarcasm, and it may get difficult to analyse the

Classification methods	Accuracy
Kernel SVM	0.6843
SVM	0.6372
RF	0.6321
Decision tree	0.603
KNN	0.562
Naïve Bayes	0.4268

Table 7.
Accuracy of models

abbreviations and acronyms. This problem can be overcome in the future by creating even more effective models.

Explainable artificial intelligence tools such as LIME, DeepLift, Shapely and What-If can be used to interpret the machine learning models well. Deep learning can be used to improve the accuracy and other classification metrics such as F1-Score, specificity, recall, precision and L-measure for different models.

Conclusion

Predicting the stock prices can be a complex and a very important activity to be performed by an investor, especially when the prices are impacted by various factors. Hence, to ease this process and make it easy for an investor to analyse the news along with fundamental analysis, classification models are formulated.

If the news is positive, then it can be said to have a bullish impact on the market, and the price of the share will rise. If the news turns out to be negative news, then it could be said that it is going to have a negative impact and result in downward price movement of the stock price in the market. And lastly, if news is neutral, then we can see that it will have no impact in the stock market and will have more stability in the market.

In this study, it can be concluded that news analysis in stock prediction gives accuracy between 42% and 68%. Kernel SVM has the highest accuracy, which is 68%, followed by SVM with 64%, then accuracy followed by RF is also considerable around 63%, then next highest accuracy is by decision tree, which is 60%, then KNN with 56% and lastly, the lowest accuracy is by naive Bayes with 42%. Hence, the model created using Kernel SVM would be the best model to help understand the market sentiment and determine the price predictions.

References

Bhardwaj, A., Narayan, Y. and Dutta, M. (2015), "Sentiment analysis for Indian stock market prediction using Sensex and nifty", *Procedia Computer Science*, Vol. 70, pp. 85-91.

Das, S., Behera, R.K. and Rath, S.K. (2018), "Real-time sentiment analysis of twitter streaming data for stock prediction", *Procedia Computer Science*, Vol. 132, pp .956-964.

Gite, S., Khatavkar, H., Kotecha, K., Srivastava, S., Maheshwari, P. and Pandey, N. (2021), "Explainable stock prices prediction from financial news articles using sentiment analysis", *PeerJ Computer Science*, Vol. 7, p. e340.

Jangid, H., Singhal, S., Shah, R.R. and Zimmermann, R. (2018), "Aspect-based financial sentiment analysis using deep learning", *Companion proceedings of the web conference 2018*, pp. 1961-1966.

Jin, Z., Yang, Y. and Liu, Y. (2020), "Stock closing price prediction based on sentiment analysis and LSTM", *Neural Computing and Applications*, Vol. 32 No. 13, pp. 9713-9729.

Jing, N., Wu, Z. and Wang, H. (2021), "A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction", *Expert Systems with Applications*, Vol. 178, p. 115019.

Joshi, S. (2016), "Financial sector development and economic growth in India: some reflections".

Kinyua, J.D., Mutigwe, C., Cushing, D.J. and Poggi, M. (2021), "An analysis of the impact of president trump's tweets on the djia and S&P 500 using machine learning and sentiment analysis", *Journal of Behavioral and Experimental Finance*, Vol. 29, p. 100447.

Li, Q., Chen, Y., Jiang, L.L., Li, P. and Chen, H. (2016), "A tensor-based information framework for predicting the stock market", *ACM Transactions on Information Systems (TOIS)*, Vol. 34 No. 2, pp. 1-30.

Mehta, P., Pandya, S. and Kotecha, K. (2021), "Harvesting social media sentiment analysis to enhance stock market prediction using deep learning", *PeerJ Computer Science*, Vol. 7, p. e476.

Mehtab, S. and Sen, J. (2019), "A robust predictive model for stock price prediction using deep learning and natural language processing", available at: SSRN 3502624.

Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015a), "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques", *Expert Systems with Applications*, Vol. 42 No. 1, pp. 259-268.

Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015b), "Predicting stock market index using fusion of machine learning techniques", *Expert Systems with Applications*, Vol. 42 No. 4, pp. 2162-2172.

Picasso, A., Merello, S., Ma, Y., Oneto, L. and Cambria, E. (2019), "Technical analysis and sentiment embeddings for market trend prediction", *Expert Systems with Applications*, Vol. 135, pp. 60-70.

Sagala, T.W., Saputri, M.S., Mahendra, R. and Budi, I. (2020), "Stock price movement prediction using technical analysis and sentiment analysis", *Proceedings of the 2020 2nd Asia Pacific Information Technology Conference*, pp. 123-127.

Shuhidan, S.M., Hamidi, S.R., Kazemian, S., Shuhidan, S.M. and Ismail, M.A. (2018), "Sentiment analysis for financial news headlines using machine learning algorithm", *International Conference on Kansei Engineering & Emotion Research*, Springer, Singapore, pp. 64-72.

Sohangir, S., Wang, D., Pomeranets, A. and Khoshgoftaar, T.M. (2018), "Big data: deep learning for financial sentiment analysis", *Journal of Big Data*, Vol. 5 No. 1, pp. 1-25.

Souma, W., Vodenska, I. and Aoyama, H. (2019), "Enhanced news sentiment analysis using deep learning methods", *Journal of Computational Social Science*, Vol. 2 No. 1, pp. 33-46.

Valle-Cruz, D., Fernandez-Cortez, V., López-Chau, A. and Sandoval-Almazán, R. (2022), "Does twitter affect stock market decisions? Financial sentiment analysis during pandemics: a comparative study of the h1n1 and the covid-19 periods", *Cognitive Computation*, Vol. 14 No. 1, pp. 372-387.

Yadav, A., Jha, C. K., Sharan, A. and Vaish, V. (2020), "Sentiment analysis of financial news using unsupervised approach", *Procedia Computer Science*, Vol. 167, pp. 589-598.

Further reading

Kalyani, J., Bharathi, P. and Jyothi, P. (2016), "Stock trend prediction using news sentiment analysis", arXiv preprint arXiv:1607.01958.

Schumaker, R.P., Zhang, Y. and Huang, C.N. (2009), "Sentiment analysis of financial news articles", *20th Annual Conference of International Information Management Association*.

Corresponding author

Gaurav Sarin can be contacted at: gaurav.sarin@iiml.org