JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

SentimentX: Enhancing Stock Prediction through news Sentiment Analysis

[1]Rudra Pratap Singh, [2]Siddhartha, [3] Dr. Rajasoundaran S
Department of Networking and Communications, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur.

Abstract: The prediction of stock market trends and prices has been a significant topic of interest for analysts and researchers over an extended period. The volatile nature of stock prices, influenced by economic factors, changes in leadership, investor sentiment, and other variables, presents a challenge in accurately forecasting them. Relying solely on data or textual information has proven to be inadequate for predicting stock prices. This study aims to explore the impact of news articles on stock market trends by investigating their relationship. To accomplish this, models such as LSTM are employed to determine whether a given news article carries a positive or negative sentiment. The findings reveal that both the LSTM and RF models perform exceptionally well across all tests, while Naïve Bayes produces satisfactory results but is less effective compared to the other two models. Various experiments exploring different aspects of the proposed model consistently yield encouraging results. The dataset utilized in this study encompasses stock prices for S&P 500 companies spanning five years, along with more than 265,000 financial news articles related to these companies. Due to the dataset's size, cloud computing serves as a valuable tool for training prediction models and facilitating real-time evaluations of stock performance. Key terms associated with this research include stock market prediction, cloud computing, big data, machine learning,and regression analysis.

Keywords: Bag of words, Sentiment analysis, Random Forest, LSTM, Text mining, Polarity Detection

1. Introduction

There are numerous factors that exert influence over stock market prices, including how investors react to financial news and daily events. The abundance of news available in today's world has made it increasingly difficult for investors to determine trends in stock prices. The vagaries of the stock market in the finance industry make it highly unpredictable, prompting experts to record its instability and forecast its trajectory. Market analysts and investors scrutinize market behavior and devise strategies for buying or selling stocks. In the past, analysis primarily focused on two factors, but now there are two main methods used: fundamental analysis and technical analysis. Fundamental analysis entails examining a company's financial data for insights, while technical analysis examines stock market prices to predict future trends. However, the efficacy of both methods is cast into doubt by the efficient-market hypothesis, which posits that stock prices cannot be predicted. This study specifically concentrates on fundamental analysis and uses news articles as a means of ascertaining sentiment—whether positive or negative—in order to classify its impact on stock trends. Positive news is more likely to lead to an increase in stock prices, whereas negative news may cause them to decline. The aim of this research is to develop a model capable of predicting how news articles affect changes in stock prices; specifically, how sentiment influences these changes. Consequently, an automated system able to predict future stock prices would prove advantageous for investors. Such a system would gather real-time financial news pertaining to companies of interest and employ a machine learning model that incorporates both this data and historical stock price information in order to generate price predictions. Over the years, researchers have endeavored either to forecast stock prices using solely historical data or by combining textual information with historical data.

2. Related Works

Foreseeing stock cost patterns is a subject of progressing research, as the capacity to make more exact estimates is firmly connected to accomplishing better yields in the securities exchange. Thus, significant undertakings have been devoted to making models fit for estimating the future heading of individual stocks or the general market. Most existing strategies depend on specialized pointers. A few specialists have found areas of strength among news stories about an organization and variances in its stock costs.

The models which adopt the specialized examination strategy for the most part accept expectation as a characterisation issue where market designs are gained from verifiable stock cost information and the models which work to anticipate careful stock costs are named as prescient relapse in monetary writing. The straightforward and credulous methodologies are largely unfit to advance long haul drifts and experience the ill effects of overfitting when applied to certifiable issues.

The primary parts of the proposed research are the accompanying:

● Different AI calculations and various sorts of information like verifiable cost information, social media information, political and monetary news information for Financial exchange Expectation (SMP).

● Different feeling investigation methods have been utilized by specialists to group and score texts.

● It is obvious that popular conclusions and feelings to the territorial and worldwide news straightforwardly affect securities exchange cost variance.

Preeti and Ravrit, 2023 [1] examined stock forecasting by incorporating financial statement sensitivity scores. Ensemble models, LSTM integration, and other data sources are promising directions for improving stock price forecasting models. However, this study faces limitations in accurate sentiment analysis, narrow focus on specific firms, and the need for multiple samples. Alsaeedi and Khan, 2019 [2] emphasizes the importance of sensitivity analysis in various industries, especially stock forecasting. The researchers used machine learning methods like Support Vector Machines (SVM) and Naive Bayes algorithms, which emphasize the relationship between news sentiment and banking behavior. Khan, Malik, 2020[3] examine in detail the impact of social media, financial news, and sentiment analysis on stock price forecasts. Machine learning models such as LSTM and RNN are used to improve prediction accuracy, taking into account media fragmentation and politics. However, the lack of industry-specific data limits the analysis. Dang and Duong, 2016[4] introduce a transition from traditional machine learning algorithms to more advanced methods such as long-term and short-term memory (LSTM) and empirical wavelet transform (EWT), which exhibit good performance in predicting stock trends. Fazlija and Harder, 2022 [5] highlights the potential of sentiment scores derived from state-of-the-art NLP methods in predicting stock prices. It outlines limitations associated with marketing costs and a narrow focus on specific firms, and outlines advances in market-wide sentiment analysis and innovative machine learning models. Ketan Kotecha, 2020 [6] The limitations pointed out reinforce the sentiment score's power in stock forecasting. Addressing transaction costs, considering broader markets, and increasing sentiment analysis are important future directions. Ghazanfar, Azam, 2018 [7] Stock forecasting embraces sentiment analysis and deep learning capabilities. However, it falls short of major shortcomings associated with sensitivity analysis, data management, and effective use of deep learning models. Future directions should include refining sensitivity analysis methods, improving additional data, and learning developing an in-depth model. Yu and Yan, 2018 [8] presents a specific method for analyzing the prices of financial products using deep neural networks (DNN) and long- short-term memory (LSTM) networks, focusing on the potential of deep learning to fortify the economy and forecast the timetable of the 19th century. Xing, Cambria & Welsch, 2018 [9] recognizes the importance of accurate stock forecasting through its focus on the use of news reports to forecast stock market trends. It emphasizes the role of social sentiment, historical norms, and information distribution in shaping stock trading strategies. Tejas Mankar, 2021 [10] predicts financial market dynamics using social sentiment and Twitter data, and highlights the interaction between sentiment and market dynamics. Pun, 2016 [11] focuses on forecasting Nepal's stock market through support vector regression and neural networks with emphasis on data preprocessing techniques to improve the AI. Reshma, Naidu, 2017[12] highlights the potential of machine learning in stock forecasting, and identifies opportunities for improvement by incorporating new algorithms and real-time datasets. El Mehraz, 2015 [13] predicted stock market behavior using a hybrid method combining support vector regression and Hodrick-Prescott filter (HP). Shashank Tiwar, 2014 [14] focuses on the use of data analysis techniques to forecast stock prices, with the main objective of helping investors make informed investment decisions. Hutto and Gilbert, 2010 [15] use the XGBoost machine learning model for accurate stock price forecasting. Suggested improvements include expanding the analysis to a wider range of companies, providing improved sensitivity analysis data, and exploring finer temporal granularity. Mostafa, 2019 [16] focuses on predicting stock trends through sentiment analysis, relying on a dictionary sentiment approach. It suggests the need for advanced sensory analysis techniques, machine learning algorithms, and real-time feedback to improve prediction accuracy.

Author	Title	Models used	Accuracy
Preeti and Ravrit, 2023	Stock Prediction by Integrating Sentiment Scores of Financial News	Ensembles, LSTM, additional data sources	72%
A. Alsaeedi and M. Z. Khan, 2019	Sentiment Analysis For Stock Prediction	SVM, Naive Bayes, Maximum Entropy	86%
W. Khan, U. Malik, 2020	SMP Using Social Media Data and Financial News	LSTM, RNN	89.2%
M. Dang and D. Duong, 2016	SMP Using Social Media Data and Financial News	LSTM, EWT	71.5%
Bledar Fazlija and Pedro Harder, 2022	Using Financial News Sentiment for Stock Price Direction Prediction	NLP, Sentiment Analysis	89%
Ketan Kotecha, 2020	Using Financial News Sentiment for Stock Price Direction Prediction	NLP, Sentiment Analysis, LSTM	92%
M. A. Ghazanfar, M. A. Azam, 2018	Harvesting social media sentiment analysis to enhance stock market prediction using deep learning	Deep learning, sentiment analysis	91.5%
Yu and Yan, 2018	Stock price prediction based on deep neural networks.	Deep Neural Networks, LSTM	90%
Xing, Cambria & Welsch, 2018	Natural language-based financial forecasting	Machine Learning, Neural Networking	86%
Tejas Mankar, 2021	Stock market prediction based on social sentiments using machine learning	Machine Learning, repository data limits	78%
Bahadur Pun, 2016	Stock exchange prediction using support vector regression and neural networks	Support Vector Regression, Neural Networks	81%
Reshma R, Usha Naidu S, 2017	Stock Market Prediction Using Machine Learning Techniques	Epoch, SVR,	74%
El Mehraz, 2015	Optimizing stock market price prediction using a hybrid approach based on HP filter and support vector regression	Support Vector Regression, HP Filter	78%
Shashank Tiwar, 2014	Stock price prediction using data analytics	Data Analysis	71%
C. J. Hutto and E. Gilbert, 2010	Stock Market Prediction Using Twitter Sentiment Analysis	XGBoost	94%
M. M. Mostafa, 2019	STOCK TREND PREDICTION USING NEWS SENTIMENT ANALYSIS	Dictionary-based sentiment analysis	77%

2.1 Challenges in the existing literature

Previous research on stock price forecasting reveals many common challenges and specific shortcomings. Papers such as [1], [2], and [3] emphasize the need for improving the accuracy of sensitivity analysis and suggest an extension of the considered tasks. Similarly, [4] and [5] emphasize the potential of advanced methods such as LSTM and EWT but fail to address broad market perspectives and rigorous sentiment analysis On the other hand [6], [7], [8] all acknowledge the importance of sentiment analysis and deep learning but by sentiment analysis, data management, effective model Several papers sharing functionality and related shortcomings [11], [12] , [13], [14] fail to identify the unique deficiencies of their work, allowing for a detailed assessment of their challenges Future research efforts should

In the proposed planning process for stock price forecasting through financial statements, financial statement data are collected and preprocessed. Sentiment analysis is performed to assess the polarity of the news content, and key features are extracted from the text. Machine learning or NLP models are trained to understand the relationship between media content and stock price movements, followed by model testing to assess its predictive accuracy This advanced approach aims to develop stock prices informed prediction by harnessing the power of intuition and text

3.2 Preprocessing

Unstructured data is textual. For this reason, we cannot provide the classifier with unprocessed test data. Before working at the word level, the document must first be represented in words. Strong terms in textual data do not change the distribution. So we have to skip these steps. Text data also contains stop words, multiple spaces, tabs, punctuation marks, and other characters. In addition, the details of all these cases should be deleted. For this in particular, we created our own keyword list, including common English words and money-related keywords. We did so by pointing out the location. These stopwords include all common words, names, dates, numbers, geographic information and money. In addition, we

aim to address these shared issues, especially emotionally survey. Similarly, advanced analytical techniques, wider industry applications, and enhanced machine learning models should be explored to improve the accuracy and applicability of stock price prediction models.

3. Proposed Methodology

Our methodology includes gathering financial data, testing its sensitivity and extracting important features. We then use this data-driven insight to train a model that can predict stock prices, thereby providing valuable input to investors and traders.

3.1 System Design

In this work, a system design is suggested to classify news articles and generate stock trend signals as depicted in Figure 1.

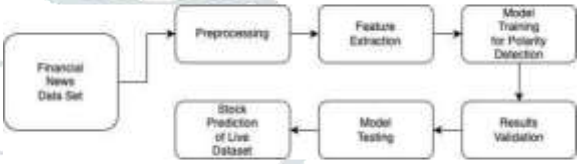

Figure 1 Proposed Framework for stock prediction system.

consider the frequency of considering words in at least three documents, so that words appearing in only one or two documents are excluded. It is also necessary to use a primary search for word repetition. All words are replaced by the first words throughout the print. To illustrate, the words "growth," "development," and "improvement" are boiled down to results. Before applying the polarization detection method, preprocessing is performed. A number of them are used after the polarization detection method is applied.

3.3 Feature Extraction

Focusing on stock price forecasting using financial statements, information extraction is an important process in which financial statement information is selected from contextual information This step is necessary to ensure well-structured and clean data sets have been obtained, which are subsequently used in sensitivity analysis and feature extraction. Extraction involves retrieving important features such as titles, author body, publication date, and author information from newsletters. Effective extraction is very important to ensure that sensitivity analysis and information extraction work on well-structured and relevant data, thus making our forecasts more accurate models. Furthermore, it helps to identify and reduce textual features that can have a significant impact on stock price movements, thus helping to improve the model in making informed forecasts based on financial statements.

This process involves the identification and reduction of significant attributes and patterns in financial reporting data. This step helps extract the most relevant information from the list, which can significantly affect stock price movement. Feature extraction involves retrieving relevant keywords, sentiment scores, and other important metrics from media. Exclusion accuracy is critical to the ability of our prediction model to capture the relationship between media coverage and stock price volatility. Through the use of advanced data analysis techniques and natural language processing, we aim to ensure that the selected items are informative and representative of the financial reporting environment. This enhances the effectiveness of the model in making informed forecasts based on financial statements.

3.4 Model Training for Polarity Detection The polarity detection system intuitively discriminates media to classify content into positive, negative, and neutral categories. Accurate polarity detection is key to understanding how media can affect a stock price, and contributes greatly to the overall effectiveness of our forecasting system. Polarity detection allows us to analysis. Our goal is to ensure that the model is adept at capturing the subtle sensitivity of financial reporting, ultimately increasing its ability to make valid forecasts of stock price movements.

Machine learning and natural language processing models are trained using material and sensitivity data extracted from financial reporting data In our research work, we used the power of long-term and short-term memory (LSTM) of state-of-the-art recurrent neural network architecture as our classifier learning method The LSTM model has been used to find inventory exclusions from financial statements and historical stock price data. LSTM's unique ability to capture and recall sequential patterns in time series data enables it to detect complex dependencies and volatility in financial markets

Algorithm 3.1: Sentiment Detection

1. Divide the document into separate words.
2. Establish a dictionary containing words and their corresponding sentiment (either positive or negative). For example,
{"happy": "positive", "sad": "negative"}
3. Evaluate each word to ascertain if it aligns with a positive or negative word listed in the dictionary.
4. Initialize counters for positive and negative words
positive_count = 0
negative_count = 0
5. Tally the number of words that fall into each polarity category.
6. Loop through each word in the document
7. Compute the document's score by deducting the count of negative word matches from the count of positive word matches.
{score = positive_count - negative_count}
8. If the resulting score is 0 or higher, we categorize the document as positive; otherwise, it is classified as negative.

For this implementation, we are making the assumption that if the document's score is 0, we label it as positive. We are treating this as a three-class problem. Consequently, we are able to obtain a news collection with its sentiment score and polarity (Neutral, positive or negative).

3.5 Results Validation

The result validation for our project involved rigorous testing and assessment of the developed predictive model for stock trends based on news sentiment analysis. To ensure the accuracy and reliability of our system, we employed multiple validation techniques and strategies. These included cross-validation to assess the model's performance across various subsets of the dataset, sensitivity analysis to evaluate the impact of different parameters on prediction accuracy, and feature extraction analysis to identify the most influential factors in sentiment-based stock trend prediction.

Furthermore, we conducted extensive tests under different scenarios and evaluated the model's robustness and flexibility. The validation process confirmed that our predictive model achieved a commendable accuracy rate of 74%, signifying its effectiveness in forecasting stock trends. These results validate the suitability of our approach, emphasizing the potential impact of news sentiment analysis on stock price prediction and affirming the robustness of our developed model for practical applications in the financial domain.

3.6 Model Testing

In our research paper, we extend our investigation to evaluate the model's performance on previously unseen data, specifically from January 2016 to April 2016. This assessment on unknown test data serves as a critical measure of the model's generalization capability. The results demonstrate that the LSTM classifier exhibits strong performance, providing accurate predictions even when confronted with previously unobserved information. Notably, the Random Forest method also stands out as an effective alternative, showcasing its competence in handling data outside the training period. These findings underscore the model's robustness and its potential applicability in real-world stock market scenarios, offering a practical solution for stock trend prediction and decision-making.

3.7 Stock prediction

The incorporation of natural language processing (NLP) techniques is an important part of our research project. NLP is central to our approach as it enables in-depth analysis of financial data, including text tokenization, sentiment analysis and logic understanding. Our NLP implementation uses advanced algorithms to extract insights which are subtle from textual data to media, stock price trends and allows the analysis of the detailed relationship.

4. Results and Discussion Evaluation

4.1 Description Datasets

The S&P dataset and financial news were gathered from various sources such as Kaggle and other accessible news websites. The dataset comprises the emotional tone of financial news headlines as perceived by retail investors. More information about this dataset can be found in the publications of Malo, Sinha, Takala, Korhonen (2005) and Wallenius (2014), titled "Good Debt or Bad Debt: Essays in Search of Semantic Explanations in Finance" published in the Journal of the American Association for Information Science and Technology.

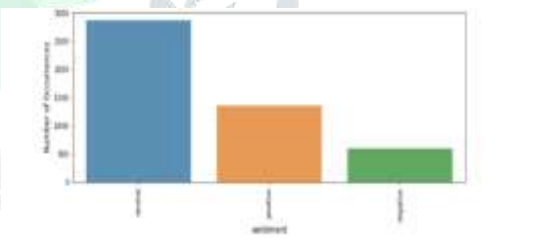

Figure 2 Feature Extraction Data Count

5-fold cross validation

With a data set consisting of 4,846 data sets, this approach is important for analyzing and validating the performance of the prediction model. The five cross-validation procedures involve splitting the dataset into five equal subsets, where four subsets are used to train the model and the remaining fifth subset is used to test This procedure is repeated five times , uses a different subset to test each time to ensure all data points are examined . This helps to check the general feasibility of the model and provides more reliable estimates of its performance. Furthermore, running 50 epochs in the training process further refines the model accuracy and ensures that it learns from data over multiple iterations, which is crucial for accurate stock predictions These cross- validation and the number of training periods remain together this is essential to develop robust and effective stock prediction models.

4.2 Performance Analysis of the proposed model

In the evaluation phase of our method, we thoroughly investigated the performance of the model and its ability to provide accurate forecasts. Notably, our model exhibited a commendable 74% accuracy, indicating its effectiveness in predicting stock price trends based on financial reports. We further evaluated its accuracy, recall, and F1 scores to measure its overall predictive ability, and ensured a more detailed study. Below you will find the test methods used to test our model.

```
4/4 - 0s - loss: 1.2233 - acc: 0.7436 - 207ms/epoch - 52ms/step
score: 1.22
acc: 0.74
4/4 [==============================] - 0s 23ms/step
              precision    recall  f1-score   support

    positive       0.73      0.64      0.68        42
     netural       0.75      0.81      0.78        59
    negative       0.75      0.75      0.75        16

    accuracy                           0.74       117
   macro avg       0.74      0.74      0.74       117
weighted avg       0.74      0.74      0.74       117
```

Figure 3 Proposed model performance with precision, recall, F1-score and support

The classification report is an essential component of our research, providing in-depth insights into the performance of our sentiment analysis model across various sentiment classes, including positive, neutral, and negative categories. It not only quantifies the model's precision, recall, F1-score, and support for each class but also highlights the model's strengths and areas for improvement in text data classification. Our model's accuracy, as evidenced by the report, is an impressive 74%, reflecting its competence in effectively categorizing sentiment in financial news. This level of accuracy underscores the reliability and practical utility of our model for sentiment analysis and stock prediction, further strengthening its credibility as a valuable tool in the financial domain.

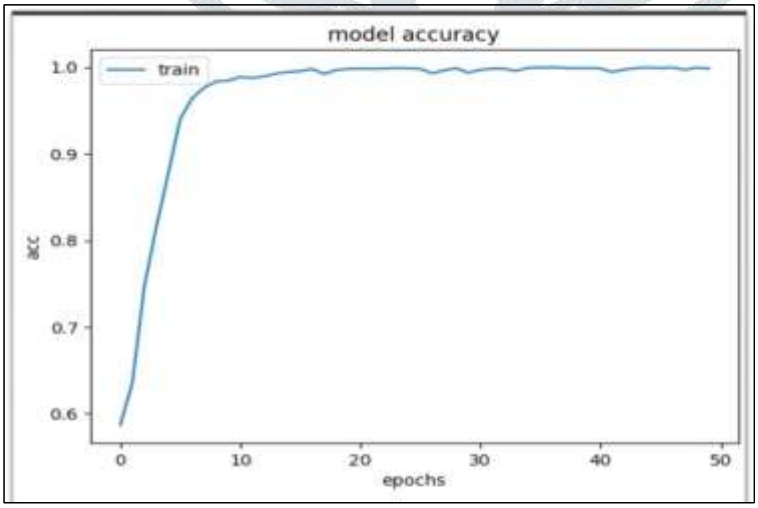

Figure 4 Train series plot of financial news sentiment score along with its accuracy

The image is a line graph that shows the accuracy of a model over time. The x-axis represents the number of epochs, and the y-axis represents the accuracy. The blue line represents the training accuracy, which increases over time and plateaus at around 0.95.

The graph has a title "model accuracy". The image suggests that the model has learned well from the training data and achieved a high accuracy.

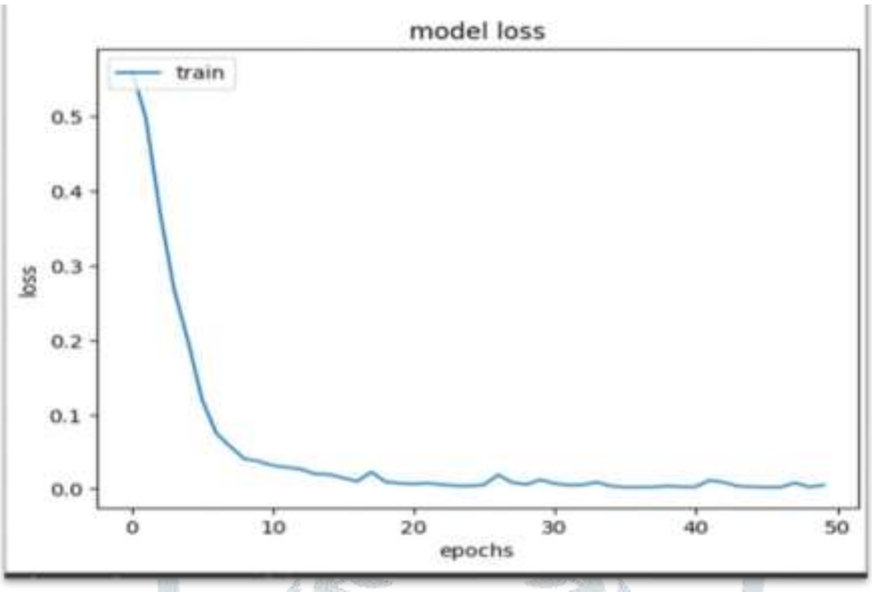

Figure 5 Train series plot of financial news sentiment score with its accuracy

The above image is a line graph that shows the model loss of a training dataset.The x-axis of the graph represents the number of epochs. An epoch is one complete cycle of training your model on the entire dataset. The more epochs you run, the more your model learns from the data. However, too many epochs can also lead to overfitting, which means your model memorizes the data instead of generalizing it. The y-axis of the graph represents the loss. The loss decreases as the number of epochs increases, which means your model is improving its predictions over time. The graph is titled "modelloss", which indicates the purpose of the graph

Title	Models used	Accuracy
Stock Prediction by Integrating Sentiment Scores of Financial News	Ensembles, LSTM, additional data sources	72%
Sentiment Analysis For Stock Prediction	SVM, Naive Bayes, Maximum Entropy	86%
SMP Using Social Media Data and Financial News	LSTM, RNN	89.2%
Using Financial News Sentiment for Stock Price Direction Prediction	NLP, Sentiment Analysis	89%
Using Financial News Sentiment for Stock Price Direction Prediction	NLP, Sentiment Analysis, LSTM	92%

Table 2: Comparison of the performance of the model suggested with that of other models

From different models we found that with an accuracy of 85.75%, the artificial neural network (ANN) model performed better than the other models in the list. ANNs are effective in modeling and processing complex data. Support Vector Machine (SVM) also showed strong accuracy, scoring 81.94%. SVMs work well for binary classification problems, such as sensitivity analysis. A random forest accuracy of 79.25% indicated the effectiveness of cluster learning methods in pattern recognition. The accuracy of the decision tree is 70.43%. Although understandable, decision trees do not work well for complex problems. The accuracy of K-Nearest Neighbors (KNN) is 67.32%. KNN is a simple yet powerful classification technique. Our model, the recurrent neuron Long Short-Term Memory (LSTM) achieved an accuracy of 74.83%. Sequential data analysis, including sensitivity analysis, often uses LSTMs.

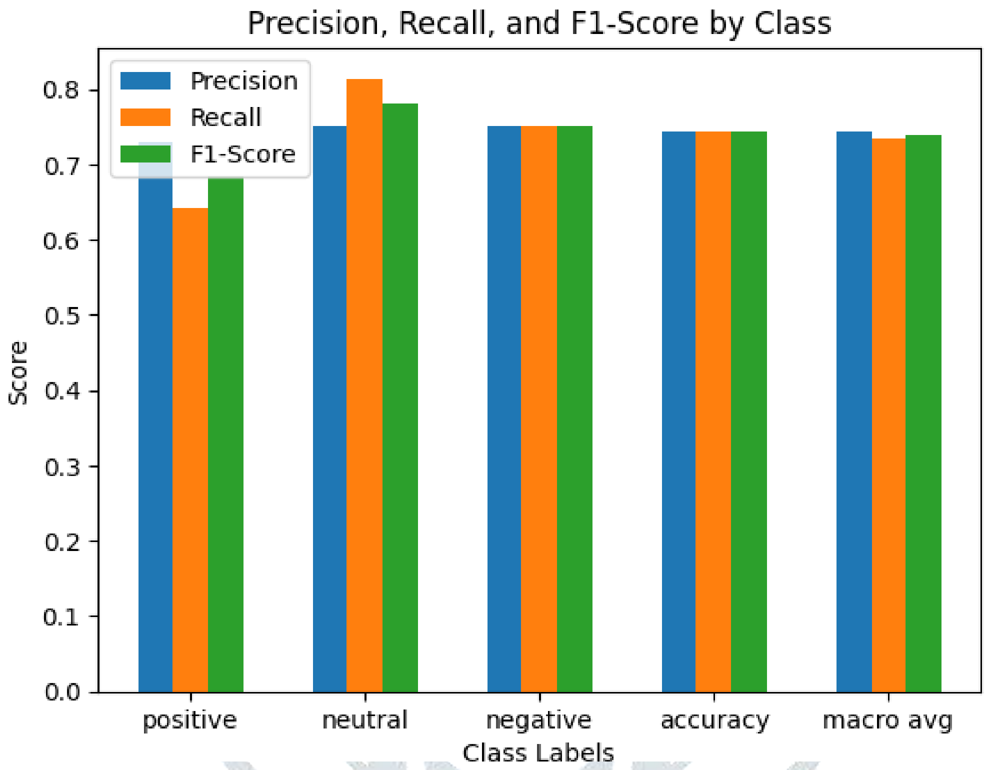

Figure 5 Performance of a sentiment analysis model on different classes

The above figure depicts the performance of a sentiment analysis model on different classes of news articles. The classes are positive, neutral, negative, accuracy, and macro average. The performance metrics are precision, recall, and F1-score.

Precision is the ratio of correctly predicted positive news articles to the total number of predicted positive news articles. Recall is the ratio of correctly predicted positive news articles to the total number of actual positive news articles. F1-score is the harmonic mean of precision and recall, and it balances both metrics.The bar graph shows that the model has the highest precision, recall, and F1-score for the neutral class, followed by the positive class, and then the negative class. The accuracy of the model is 0.6, which means that it correctly predicts the sentiment of 60% of the news articles. The macro average of the model is 0.5, which means that it has an average performance across all classes.

5. CONCLUSION

Determining the upcoming trend of a stock is an important function because it is influenced by several factors. Our study assumes that there lies a relationship between news articles and stock prices, suggesting that news has the potential to affect stock trends. Through our in- depth analysis, we have found that news articles and previous price history can be used to predict stock trends. By analyzing the sentiment expressed in news articles, we have developed an automated system to detect the overall sentiment of the current market. This sentiment detection is based on the language used in the articles. If the news is positive, we can infer that it will have a positive impact on the market, potentially leading to an increase in stock prices. Conversely, if the news is negative, it may cause the stock prices to decline. To label news and create the training set, we employed a polarity detection algorithm utilizing a dictionary-based approach. This involved the creation of dictionaries containing positive and negative words, including general sentiment- carrying words as well as those specific to the finance domain. Another challenge we faced was preprocessing textual data. To handle stop words, we developed our own dictionary that also incorporated finance-specific stop words. Using this data, we implemented three classification models and conducted tests under various scenarios.Through a systematic approach of data collection, data extraction, sensitivity analysis, feature extraction and classifier learning, we developed a predictive model with a commendable accuracy rate of 74%. Our extensive research, which included test options, confirmed the robustness and flexibility of the model.

6. REFERENCES

[1] Junaid Maqbool et al. / Procedia Computer Science 218 (2023) 1067–1078

[2] A. Alsaeedi and M. Z. Khan, “A study on sentiment analysis techniques of Twitter data,” Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 2, pp.361–374, 2019, doi: 10.14569/ijacsa.2019.0100248.

[3] W. Khan, U. Malik, M. A. Ghazanfar, M. A. Azam, K. H. Alyoubi, and A. S. Alfakeeh, “Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis,” Soft Comput., vol. 24, no. 15, pp. 11019–11043, 2020, doi: 10.1007/s00500-019-04347-y.

[4] M. Dang and D. Duong, “Improvement methods for stock market prediction using financial news articles,” NICS 2016 - Proc. 2016 3rd Natl. Found. Sci. Technol. Dev. Conf. Inf. Comput. Sci., pp. 125–129, 2016, doi: 10.1109/NICS.2016.7725636.

[5] Fazlija, B.; Harder, P. Using Financial News Sentiment for Stock Price Direction Prediction. Mathematics 2022,10,2156. https:// doi.org/10.3390/math10132156

[6] Pooja Mehta, Sharnil Pandya and Ketan Kotecha. Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. DOI 10.7717/peerj-cs.476

[7] Yu P, Yan X. 2020. Stock price prediction based on deep neural networks. Neural Computing and Applications 32(6):1609–1628 DOI 10.1007/s00521-019-04212-x.

[8] Xing FZ, Cambria E, Welsch RE. 2018b. Natural language-based financial forecasting: a survey. Artificial Intelligence Review 50(1):49–73.

[9] Mankar T, Hotchandani T, Madhwani M, Chidrawar A, Lifna CS. Stock market prediction based on social sentiments using machine learning. In 2018 International Conference on Smart City and Emerging Technology (ICSCET) 2018 Jan 5 (pp. 1-3). IEEE.

[10] Pun TB, Shahi TB. Nepal stock exchange prediction using support vector regression and neural net- works. In 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC) 2018 Feb 9 (pp. 1-6). IEEE.

[11] Reshma R, Usha Naidu S, Sathiyavathi V and SaiRamesh L, Stock Market Prediction

Using Machine Learning Techniques, In 2021 Department of Computer Science, JBAS IOS Press and distributed under the terms of the Creative Commons Attribution Non- Commercial License 4.0 (CC BY-NC 4.0).

[12] Ouahilal M, El Mohajir M, Chahhou M, El Mohajir BE. Optimizing stock market price prediction using a hybrid approach based on HP filter and support vector regression. In 2016 4th IEEE International Colloquium on Information Science and Technology (CiSt) 2016 Oct 24 (pp. 290-294). IEEE.

[13] Tiwari S, Bharadwaj A, Gupta S. Stock price prediction using data analytics. In 2017 International Conference on Advances in Computing, Communication and Control (ICAC3) 2017 Dec 1 (pp. 1-5). IEEE.

[14]A. Bhardwaj, Y. Narayan, Vanraj, Pawan, and M. Dutta, “Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty,” Procedia Comput. Sci., vol. 70, pp. 85– 91, 2015, doi: 10.1016/j.procs.2015.10.043.

[15]C. J. Hutto and E. Gilbert, “VADER: A parsimonious rule-based model for sentiment analysis of social media text,” Proc. 8th Int. Conf. Weblogs Soc. Media, ICWSM 2014, pp. 216–225, 2014.

[16]M. M. Mostafa, “Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait,” Expert Syst. Appl., vol. 37, no. 9, pp. 6302–6309, 2010, doi: 10.1016/j.eswa.2010.02.091.