

Predict Market fluctuations Based on the Twitter sentiment index (TSI) and the Sentiment of Financial Video News Sites

1st Faten Alzazah, 2nd Dr. Xiaochun Cheng, 3rd Prof Xiaohong Gao

Department of Computer Science

Middlesex University, London

FA974@live.mdx.ac.uk, X.Cheng@mdx.ac.uk, X.Gao@mdx.ac.uk

Abstract—Scientists have long been interested in forecasting stock market fluctuations. Traditional data like financial textual news, stock prices, and comments are simply no longer sufficient because they don't provide a comprehensive picture. Some of those responses may be unfavorable. You must fully understand what they are affirming before you can determine what to do.

In this study, the efficacy of using financial video news stories versus the use of conventional text news stories to forecast the stock market is examined. We used the Granger causality test to evaluate the robustness of the causal connection between share prices, text news sentiment, video news sentiments, and the Twitter sentiment index.

Several models for sentiment analysis of S&P 500 stock were assessed using LR, SVM, LSTM, ATT-LSTM, and CNN models. This study is distinctive because it compares the use of financial video news stories, conventional text news stories, and Twitter sentiment index to forecast stock market movements. The experimental results suggest that there is a stronger causal connection between video news sentiment and stock market fluctuation compared to conventional text news sentiments. The result shows that we can more accurately predict market changes using video news than we can with traditional news.

Keywords—Market prediction, text news, video news, machine learning, Twitter sentiment index (TSI).

I. INTRODUCTION

The main objective of every investor is to forecast future stock market behavior. Currently, given the abundance of market data sources available, machine learning has become more popular than before for predicting stock market prices. Machine learning techniques enable the speedy analysis of more complicated heterogeneous data to produce more precise findings [1]. A variety of algorithmic techniques, including logistic regression (LR), Support Vector Machine (SVM), and Artificial Neural Networks (ANN), have been successfully applied to stock market prediction [2]-[4]. A text classification approach called sentiment analysis is exploited to categorize recorded thoughts or expressions into categories like positive, positive, and neutral [5]. The two main categories for stock forecasting depending on the

information sources each approach uses are fundamental analysis and technical analysis.

Technical analysis examines trading activity, such as previous prices to identify statistical trends that are used to analyze companies. Given that the needed data is consistently released [6], technical analysis is frequently employed for short-term predictions. By computing technical patterns from historical prices and volumes, technical analysis derives useful information about stocks that are then used as input to a forecasting model. Typically, the prediction of stock prices is highly erratic and unstable [7]. Therefore, relying solely on technical analysis and numerical price information to identify volatile market patterns is insufficient. Fundamental analysis, on contrary, incorporates information from many other data sources about the stock market, such as news, social media, and business reports, as added inputs for stock prediction models [8]. Prior studies have disregarded a crucial informational resource: video news, which is frequently presented by industry professionals. This study is seen as an expansion of our earlier work [9]. Compared to our previous work, the usefulness of using longer period data to predict stock price was examined in this study. We also compared the effectiveness of using video news stories versus the use of usual news text stories. We used Granger causality analysis to investigate the strength of the relationship between stock market prices and video news sentiments, usual financial text news stories, and the Twitter Sentiment Index (return). We input our ML classification models with video news sentiments and usual news text stories to find if this would improve the model's accuracy. The remaining portions of this work are outlined as follows: Section II addresses machine learning models used in this study. Section III explains the methods used in this research. In section VI, we present the outcome. Finally, section VII concludes our research.

II. Machine learning models

This section gives a general overview of machine learning models used in this research, such as Logistic Regression, Support

vector machines(SVM), Convolutional Neural Networks (CNN), Long short term memory (LSTM), Attention LSTM (Att-LSTM).

2.1 Logistic Regression

A commonly used statistical modeling approach is logistic regression. The model is simple to train and produces results with insightful predictive analysis. Linear regression is essentially the first forecasting model, For example, a line that better fits the sets of data in the data source. As a result of the prediction, any new point will be placed on such a line [10]. The model finally reveals the optimal 0 that minimizes the related cost function, resulting in a partitioning sigmoid curve among the two classifications [11]. To prevent overfitting, the technique can be normalized, but it is generally restricted to linear independent data. Text categorization and sentiment analysis benefit from selective classification algorithms such as logistic regression. The model evaluates two classes in the training phase and picks the most suitable separation [12].

2.2 Support Vector Machines (SVMs)

SVMs have supervised machine learning models that are widely used in regression and categorization applications. SVM is a hyperplane that uses a maximum margin to split a set of instances into several distinct groups [13]. Using the vectors technique, the model is used to classify texts. It seeks the best margin hyperplane for dividing the upward and downward trend classes so that the margin among these classes is as wide as possible. SVM has demonstrated recovery to the over fitting problem, accomplishing satisfactory generalization performance.

SVM can classify new texts when provided with any collection of labeled training data for distinct classes. When handed limited data, it could operate to its highest capacity [14]. Particularly in comparison to ANN, SVM is also faster and more effective [15]. The main objective of SVM in video text sentiment analysis is to find the maximum margin hyperplane for separating a series of different video texts into positive and negative group sentiments.

2.3 Convolutional Neural Network (CNN)

Authors in [16] introduced the use of Convolutional Neural Networks (CNN) for natural language processing (NLP). A traditional Convolutional neural network is comprised of several Convolutional layers at the bottom of a classification model. Text processing inputs are typical letters, words, lines of text, or transcripts that are transformed into a similarity matrix representation. Each matrix row illustrated a token, which is commonly a term or character [17]. The CNN classifier's main benefit is that it is focused on multi-layer network infrastructure and convolution layout. Convolutional layers of CNNs can construct various semantic pieces of sentences and discover the relationships among structured fragments because they apply a convolution procedure to the input matrix, allowing them to fully utilize the inter-modal semantic relatedness of market news [18]. The outputs of the CNN model denote the fluctuation of daily stock markets. The structure of text CNN is implemented by a Python library, TensorFlow.

2.4 Long Short Term Memory (LSTM)

It is a regular recurrent neural network that eliminates the exploding gradients issue. Using a memory unit and a gate technique to control how data stored in the memory cell can be utilized and updated, LSTM can represent the lengthy dependencies in a series [19]. The LSTM neural network would learn and decide whether a specific output should be the following recursive input. Every LSTM is a collection of cells or system components that capture and hold data flows. The cells form a transport line that transports data from one unit to the next and captures it for the current unit. Data can be discarded, filtered, or introduced for the next cells by using specific gates in every cell [20]. Each memory cell in LSTM has three gates that preserve and alter its cell state s_t : a forget gate (f_t), an input gate (i_t), and an output gate (o_t). The equations below are an update of the LSTM layer memory cells at each timestep t [21].

$$f_t = \text{sigmoid}(W_{f,x} x_t + W_{f,h} h_{t-1} + b_f). \quad (1)$$

$$\tilde{s}_t = \tanh(W_{\tilde{s},x} x_t + W_{\tilde{s},h} h_{t-1} + b_{\tilde{s}}). \quad (2)$$

$$i_t = \text{sigmoid}(W_{i,x} x_t + W_{i,h} h_{t-1} + b_i). \quad (3)$$

$$s_t = f_t \odot s_{t-1} + i_t \odot \tilde{s}_t. \quad (4)$$

$$o_t = \text{sigmoid}(W_{o,x} x_t + W_{o,h} h_{t-1} + b_o). \quad (5)$$

$$h_t = o_t \odot \tanh(s_t). \quad (6)$$

Where x_t denotes the input vector at time step t . $W_{f,x}$, $W_{f,h}$, $W_{\tilde{s},x}$, $W_{\tilde{s},h}$, $W_{i,x}$, $W_{i,h}$, $W_{o,x}$, and $W_{o,h}$ are weight matrices. b_f , $b_{\tilde{s}}$, b_i , and b_o are vectors of bias. s_t and $\tilde{s}_t$ are cell state vectors. Candidate values f_t , i_t , and o_t are vectors for the expert gates activation values. h_t is a vector for the LSTM layer output. After the last part of the series has been processed, the final output for the entire series is handed back.

LSTM components are intended to improve RNNs' memory for earlier values and give learning from current data priority over learning from historic data in sequential order [22] for building a prediction model for the current study, this method offers a useful foundation when needed to preserve vital data.

2.5 Attention LSTM (Att-LSTM)

Sentiment analysis is proposed to utilize the attention LSTM model [23]. Att-LSTM Offered a design process that could identify the critical section in a text that is associated with a certain feature [24]. H_i (hidden state memory vectors) are used for aspect-embedded attention architecture to produce an attention vector α , which is employed to construct a weighted or sorted presentation of H_i . The sentence expression S_i of the i th aspect was then created using the weighted notation. This permits the attention model to concentrate on the appropriate sentence aspect sections [25]. In the domain of sentiment analysis, the attention model outperforms the conventional LSTM model in terms of classification performance.

III. METHODOLOGY

We introduce a machine learning approach consisting of the process steps presented in figure 1 to predict stock market volatility. In the first step, we gathered three types of stock market data from online sites. The dataset represents (i) past close

price data for S&P 500 equities published on Yahoo Finance website. ; and (ii) A collection of video news stories regarding the S&P 500 stock from various news organizations, like CNBC, Bloomberg, YouTube, Yahoo Finance, and FX EMPIRE. iii) The 200 S&P 500 components with the most optimistic sentiment on Twitter make up the Twitter Sentiment Index, which is intended to reflect the achievement of those 200 companies.

Each of these datasets spans the period from 1/2/2019 to 12/31/2021. The speaker's voice is amplified and made louder in steps 2's conversion of video news to text news, and the recording is done using the Google Docs recording tool. However, if the video news was collected from YouTube then we copy the video transcript that is shown below the video progress bar.

In step 3, we perform text pre-processing, and then, utilizing NLP techniques, we extract numerous valuable characteristics from textual news. In step 4, We employ the Python (2 and 3) library package for text data processing called TextBlob as our instrument for polarity analysis. This method provides a simple API for delving into common natural language processing (NLP) processes for categorizing news, after which each news item is given a range between 0 and 1. Where 1 is the positive polarity with the highest polarity. In step 5, to ascertain whether there is statistically meaningful causation between the sentiments impacted by each video news item, text news item, and tweeter sentiment index and changes in stock prices, Granger Causation is used. In step 6, Applying TextBlob, a sentiment score is assigned to each text news item and each video news item. We assume that the mood and information in video news represent the market expert's individual opinions on the selected firm stock. In step 7, we find the correlation score of video news items, text news items, and tweeter sentiment index with stock price data, to decide what types of data give us the best result. Eventually, utilizing our unique dataset, we applied ML machine learning algorithms to evaluate our model accuracy. Figure 1 depicts our strategy, which is detailed in the next section.

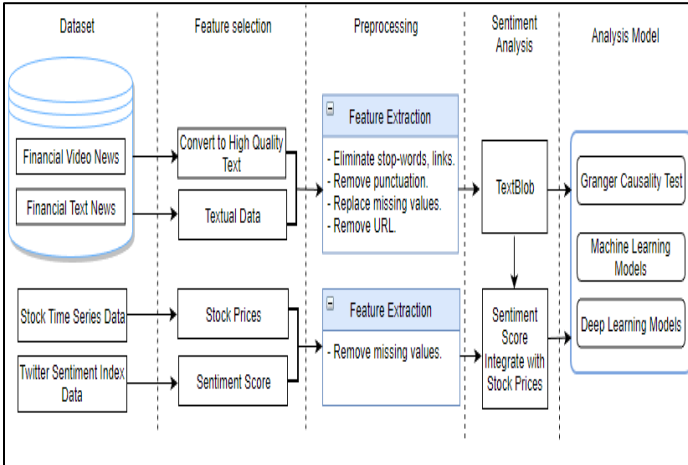


Fig. 1: The design of our suggested approach.

IV. Dataset Collecting And Text Preprocessing

Using TextBlob, we prepare the input text for classification as either positive or negative text through the news extraction and preparation step. This process is separated into various parts: 1. Seeking S&P 500-related business video and text news for the relevant period 1/2/2019 to 12/31/2021 excluding the holiday period. We have collected 726 video news and 726 text news

which came from a variety of news sites, including Bloomberg, CNBC, YouTube, Yahoo Finance, and FX EMPIRE. 2. To make the financial news videos louder and clearer for recording, we play them while they're connected to a wireless microphone. 3. If the video news is collected from YouTube then we copy the video transcript that is shown below the video progress bar. 4. If the video is not collected from the YouTube website then we use the Google docs recording feature to record the audio and add it to a text document. 5. Through using Natural Language Toolkit framework NLTK, we performed data cleaning for our text during pre-processing, which makes text easier to analyze. The scheme is a Python NLP library. At this stage, we eliminate stopwords, links, punctuation, and other parts. We also tokenize words, which involves breaking the text up into individual words or phrases. This makes it possible for us to work with shorter passages of text that are still fairly coherent and understandable when read separately from the rest of the text.

In the Final stage, Glove, one of the most popular algorithms for creating a representation of words, allows for quick training of large corpora and is also one of the most efficient ways to create word representations; therefore we prepare word embeddings for each word using it. This stage results in a vector representation of each word. The final dataset is illustrated in figure 2 and is composed of positive sentiment and negative sentiment for each of TIS, price data, video news, and text news. As the chart shows about half of the data in both price data and TIS data represent negative news. While a very small amount of data represent negative sentiment in video news and text news. The price sentiment is scored based on the fact that if the price data is higher than the previous day's price data then the sentiment will be positive and vice versa.

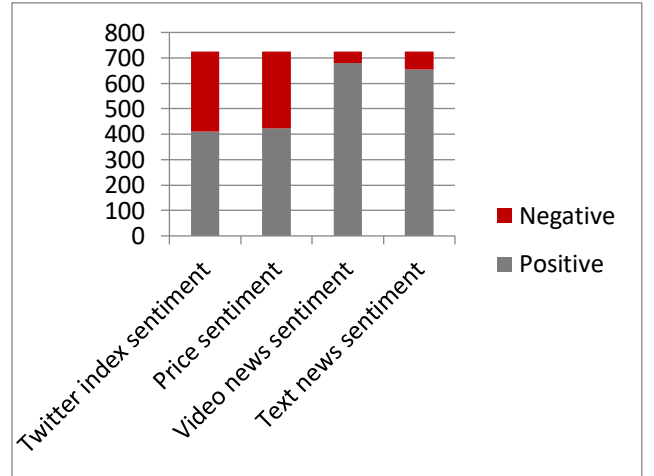


Fig. 2 Positive and negative sentiment distribution in our dataset.

V. Stock price fluctuations with News Sentiment and TSI

Examining whether there is a strong connection between the sentiment analysis of financial video news, conventional text news, and tweeter sentiment index and market fluctuations is one of the aims of the study. We conduct the main investigation into the association between stock market sentiment and news sentiment using the Granger causality analyses. Several studies have typically found a connection between market reaction and opinions expressed in either financial news sites or social media [26][27]. The Granger-causality method is utilized in this study to examine the causality relationship between the sentiments of

video news stories, conventional text news, tweeter sentiment index, and S&P 500 stock price fluctuation. Determining the period to be used in associating the release of video news and stock market changes is the first step in our investigation.

In our research, the stock market is investigated on the same day as the news is evaluated over some time. The daily timeframe could be favorable for day traders since they carry out their investments on the same day. Then, we examine textual data using the TextBlob polarity analysis tool. A score between 0 and 1 is awarded to each news article. Where 1 is the highest positive polarity score. The association between news and the financial market prices was then examined using the Granger Causality test on the stock markets' daily closing prices past data.

A. Granger Causality Test

Granger (1969) was the one who initially established the Granger Causality Hypothesis Test [28]. The hypothesis aimed to determine the short-term interactions and causal relationships between specific historical occurrences. In the practice, the exam evaluates a user's potential to forecast future data in time series using past data. In practice, the exam evaluates a user's potential to forecast future data in time series using past data. According to the Granger test, if two variables are connected then a bidirectional or unidirectional causation relationship must occur. The method's goal is to establish the causal relationship between every variable. It is claimed that there is no relationship between the variables if the time series observations are not stationary at $I(0)$. We investigate to find out whether the sentiment expressed in video news, text news and tweeter sentiment index is related to changes in the stock market, accounting for the closing prices of S&P 500 stocks. We used S&P 500 stock for sour study because it is a reliable indicator of market activity in the US. To achieve this, we tested news sentiment and stock price time series using the Granger causality analysis.

Table 1: Granger Causality test result

Lag number	TNews & price p-value	VNews price p-value	& Price & TSI p-value
1	0.0069	0.0022	0.3361
2	0.0258	0.0031	0.0112
3	0.0373	0.0124	0.0086

The results in Table 1 show that, for nearly the entire Lag period of our data, the p-value is significantly less than the 5% significant value, enabling us to reject the null hypothesis. After determining the Granger correlation between the sentiment polarity of text and video news and actual S&P 500 stock prices. We also, test the link of returns between the recently completed S&P 500 Twitter Sentiment Index and stock close price. According to the Granger causality study, there is a statistically significant causal relationship between stock prices and sentiments produced by either video news, text news, and Twitter Sentiment Index when all are studied across several lag periods. However, comparing the p-value of text news with the p-value of video news. The association between market movements and

video news is stronger than that with text news, as evidenced by the fact that the p-value of video news was generally lower than that of text news. As a consequence, since VNews offer reliable predictions of stock closing prices, sentiment features gathered from VNews can be used more effectively in our model. Also, as the table shows TSI p-value was lower than 5% for the second and third lag periods which indicated that we can also use TSL as new feature when we build our predictive model.

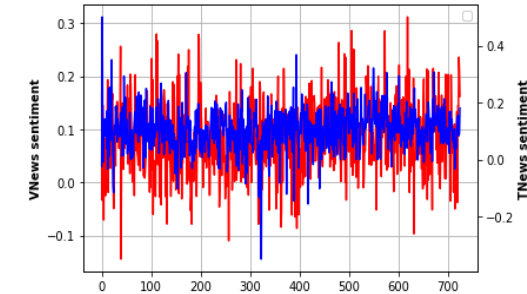


Fig.5:Granger Causality of VNews and TNews.

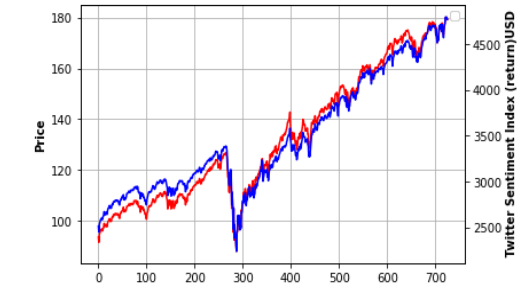


Fig.4 Granger Causality test between TSI and price.

In general, since sentiment features offer forecasting information concerning stock closing prices, they will be employed in our model. For instance, we found that just the entire lag period was statistically useful for the prediction purpose when using the VNews. Figures 4-5 illustrate the relationship between VNews, TNews, the sentiment index of Twitter, and the price movements.

B. Training of Different Machine Learning Models

Sentiment analysis of the S&P 500 stock was analyzed using several models, including LR, SVM, LSTM, ATT-LSTM, and CNN. The identified data set was separated into two segments: 90 % for training and 10 % for testing. Every model is trained using a collection of 726 video news clips. The Textblob library has labeled all of the financial video news.

In particular, news that positively affects the financial market was designated as positive, whilst news that negatively affects the market was designated as negative. Collecting texts from video news was preprocessed before each of these machine-learning techniques was employed. The performance of the five machine learning algorithms was then tested using our brand-new dataset of S&P 500 stocks in a set of experiments. The CNN connection from the Tensorflow Python module is used in this work to create the CNN deep learning model. The CNN convolution layers efficiently build different semantic segments of sentences and identify the links among generated portions when they apply a Convolutional process to the input vector. This enables them to

completely exploit the internal model semantic linkages of video news. To break down a word into tokens or phrases for deep learning models we'll initially use the Tokenizer class from the NLTK Python package. We will then construct a feature matrix using GloVe word embeddings, utilizing terms as keys and their associated embedding lists as values. The densely connected layer is then linked to our embedding layer. Next, a dense layer of sigmoid activation is added. With accuracy measures and binary cross-entropy loss function, we built our model using the Adam optimizer.

To extract features from our data for the CNN model, we used a 1D classifier model, considering that text data has only one dimension. We built a simple CNN model with a single convolutional layer. Following that, we construct our model with 128 filters (kernels) and a kernel size of 5, and a pooling layer. A sigmoid was used for the activation function. Upon that, we add a dense layer using sigmoid activation. We first create the embedding layer before initializing a sequential model to produce the LSTM model. Afterward, we construct a 2-neuron LSTM layer with 6 epochs and a batch size of 20. Using attention LSTM, To concentrate on the critical data, the attention technique ensures that the result is not the prior result of all hidden layers but rather the weighted combination of hidden layers. Also, we have used PyTorch which offers important utility functions like pad sequences. Masking and padding are addressed by such functions, so once a sentence ends, the resultant word representations are just zeros.

To predict future market movements, the attention technique can incorporate the news sentiment trend. For both SVM and LR, after textual pre-processing, we used the Term Frequency Inverse Document Frequency model (TFIDF) to convert textual data into a matrix of TFIDF features. Subsequently, to contrast the outcomes of the various machine learning methods, we supplied our data to SVM and LR models in the same manner. Section 4.4 describes the results from this phase and illustrates the ideal model for our dataset.

VI. Result and Discussion

This section describes the efficiency of our recommended machine-learning techniques, which include LSTM, ATT-LSTM, CNN, LR, and SVM.

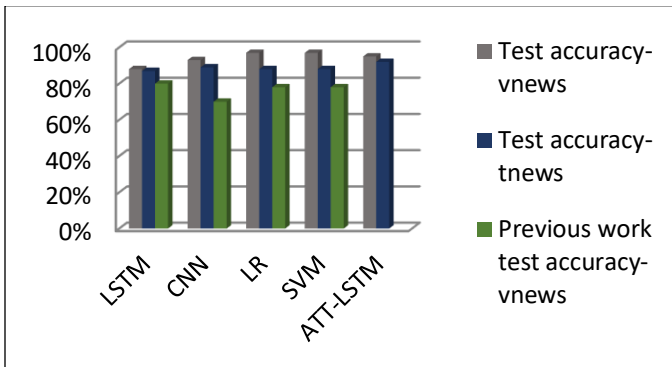


Fig 6. Accuracy of machine learning models of previous and current work. A range of statistical techniques can be used to evaluate the effectiveness of machine learning models. Equation 1 shows the accuracy measures used in this study, and the reference [29] provides more information.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{FP} + \text{TP} + \text{TN} + \text{FN}} \quad (1)$$

Where FN stands for false refused (false negative), TP for successful identification (true positive), TN for true negative (properly refused), and FP for wrongly identified (false positive). From Figure 6, it's fairly obvious that when using the video news dataset we get higher test accuracy than when using the text news dataset technique for all machine learning models, testing accuracy achieves above 88 % in all machine learning models when using our novel dataset. It is clear that VNews achieved higher accuracy results in all machine learning models compared to our previous work.

It is also obvious that the LR and SVM perform better in testing accuracy than the deep learning technique, which reaches above 96% using VNews. However, when using text news deep learning achieves better accuracy performance compared to machine learning models. In general, the test accuracy of VNews was higher than the test accuracy of TNews in all models. This proves that adding the VNews dataset to our future prediction model will improve the prediction accuracy further. However, by gathering more datasets for training, employing a balanced dataset, and including optimization machine learning approaches, prior findings can be enhanced even further.

VII. CONCLUSION AND FUTURE WORK

Due to market instability, predicting stock market behavior is one of the most challenging tasks. As a result, the subject is of concern to both academics and investors everywhere.

Traditional data, such as stock prices, views, and economic text news, are just insufficient today because they don't present a fuller understanding. There's a chance that some of these responses won't be positive. To comprehend the complex relationship between sentiment and the volatility of financial markets, much research has been done. The effectiveness of using professional video news stories versus traditional text news stories to anticipate the stock market is compared in this study. Also, developing forecasting models based on a specific sentiment analysis without determining the best algorithm strategy to do such sentiment classification may result in false conclusions. Therefore, several models for sentiment analysis of S&P 500 stock were examined using LR, SVM, LSTM, ATT-LSTM, and CNN models.

The experimental findings demonstrate that we can forecast market fluctuations more precisely by utilizing video news than we can with traditional news. When applying SVM with our video news, the maximum classification accuracy was additionally attained. However, our technique of data collection has several drawbacks. For example, employing a wireless microphone with a slow internet connection could cause interruptions in playing videos, which will result in inaccurate text being connected with them.

To overcome this limitation, two procedures have been taken. Firstly, a part of our dataset was manually collected with the assistance of a wireless microphone, and care was taken when converting any video to a nearly completed written text. Secondly, when the video news is collected from YouTube then we copy the video transcript that is shown below the video progress bar which is more accurate than the first method.

However, this strategy is impractical if someone tries to use it to collect data over a lengthy time. By gathering more data for training, employing a more balanced dataset, and utilizing machine learning optimization techniques, prior findings may be enhanced further. We publish our updated dataset to inspire another research in the link <https://www.kaggle.com/datasets/fatensubhi/video-and-text-news-dataset>.

REFERENCES

- [1] Rouf N, Malik MB, Arif T, Sharma S, Singh S, Aich S, Kim HC. Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions. *Electronics*. 2021 Nov 8;10(21):2717.
- [2] Ali SS, Mubeen M, Lal I, Hussain A. Prediction of stock performance by using logistic regression model: evidence from Pakistan Stock Exchange (PSX). *Asian Journal of Empirical Research*. 2018 Jul 6;8(7):247-58.
- [3] Nti IK, Adekoya AF, Weyori BA. Efficient stock-market prediction using ensemble support vector machine. *Open Computer Science*. 2020 Jan 1;10(1):153-63.
- [4] Safari A, Ghavifekr AA. International stock index prediction using artificial neural network (ANN) and Python programming. In *2021 7th International Conference on Control, Instrumentation and Automation (ICCIA)* 2021 Feb 23 (pp. 1-7). IEEE.
- [5] Ain QT, Ali M, Riaz A, Noureen A, Kamran M, Hayat B, Rehman A. Sentiment analysis using deep learning techniques: a review. *International Journal of Advanced Computer Science and Applications*. 2017;8(6).
- [6] Haq AU, Zeb A, Lei Z, Zhang D. Forecasting daily stock trend using multi-filter feature selection and deep learning. *Expert Systems with Applications*. 2021 Apr 15;168:114444.
- [7] Tuarob S, Wettayakorn P, Phetchai P, Traivijitkhun S, Lim S, Noraset T, Thapisutikul T. DAViS: a unified solution for data collection, analyzation, and visualization in real-time stock market prediction. *Financial Innovation*. 2021 Dec;7(1):1-32.
- [8] Alzazah FS, Cheng X. Recent advances in stock market prediction using text mining: A survey. *E-Business-Higher Education and Intelligence Applications*. 2020 Jun 1.
- [9] Alzazah F, Cheng X, Gao X. Predict Market Movements Based on the Sentiment of Financial Video News Sites. In *2022 IEEE 16th International Conference on Semantic Computing (ICSC)* 2022 Jan 26 (pp. 103-110). IEEE.
- [10] Mokhtari S, Yen KK, Liu J. Effectiveness of artificial intelligence in stock market prediction based on machine learning. *arXiv preprint arXiv:2107.01031*. 2021 Jun 30.
- [11] Acharya A, Crawford N, Maduabum M. A Nation Divided: Classifying Presidential Speeches.
- [12] Colianni S, Rosales S, Signorotti M. Algorithmic trading of cryptocurrency based on Twitter sentiment analysis. *CS229 Project*. 2015;1(5):1-4.
- [13] Dumais S, Platt J, Heckerman D, Sahami M. Inductive learning algorithms and representations for text categorization. In *Proceedings of the seventh international conference on Information and knowledge management* 1998 Nov 1 (pp. 148-155).
- [14] Chhajer P, Shah M, Kshirsagar A. The applications of artificial neural networks, support vector machines, and long-short term memory for stock market prediction. *Decision Analytics Journal*. 2022 Mar 1;2:100015.
- [15] Vishwanathan SV, Murty MN. SSVM: a simple SVM algorithm. In *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)* 2002 May 12 (Vol. 3, pp. 2393-2398). IEEE.
- [16] Collobert R, Weston J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning* 2008 Jul 5 (pp. 160-167).
- [17] Ho CC, Baharim KN, Fatan AA, Alias MS. Deep neural networks for text: A review. In *The 6th International Conference on Computer Science and Computational Mathematics*. Langkawi, Malaysia 2017.
- [18] Li X, Shang W, Wang S. Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting*. 2019 Oct 1;35(4):1548-60.
- [19] Rao G, Huang W, Feng Z, Cong Q. LSTM with sentence representations for document-level sentiment classification. *Neurocomputing*. 2018 Sep 25;308:49-57.
- [20] Siami-Namini S, Namin AS. Forecasting economics and financial time series: ARIMA vs. LSTM. *arXiv preprint arXiv:1803.06386*. 2018 Mar 16.
- [21] Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*. 2018 Oct 16;270(2):654-69.
- [22] Eapen J, Bein D, Verma A. Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction. In *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)* 2019 Jan 7 (pp. 0264-0270). IEEE.
- [23] Long Y, Xiang R, Lu Q, Huang CR, Li M. Improving attention model based on cognition grounded data for sentiment analysis. *IEEE transactions on affective computing*. 2019 Mar 4;12(4):900-12.
- [24] Wang Y, Huang M, Zhu X, Zhao L. Attention-based LSTM for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing* 2016 Nov (pp. 606-615).
- [25] Hazarika D, Poria S, Vij P, Krishnamurthy G, Cambria E, Zimmermann R. Modeling inter-aspect dependencies for aspect-based sentiment analysis. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)* 2018 Jun (pp. 266-270).
- [26] Ranco G, Aleksovski D, Caldarelli G, Grčar M, Mozetič I. The effects of Twitter sentiment on stock price returns. *PloS one*. 2015 Sep 21;10(9):e0138441.
- [27] Ding Z, Liu Z, Zhang Y, Long R. The contagion effect of international crude oil price fluctuations on Chinese stock market investor sentiment. *Applied energy*. 2017 Feb 1;187:27-36.
- [28] Granger CW. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*. 1969 Aug 1:424-38.
- [29] Nti IK, Adekoya AF, Weyori BA. A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*. 2020 Apr;53(4):3007-57.