



# Novel Sentiment Analysis from Twitter for Stock Change Prediction

Yang Cui<sup>1(✉)</sup>, Yucen Jiang<sup>2</sup>, and Haisong Gu<sup>1</sup>

<sup>1</sup> VisionX LLC., San Jose, USA  
forcuiyang@gmail.com

<sup>2</sup> Birmingham Business School, University of Birmingham, Birmingham, UK

**Abstract.** Literature in behavioral economics and socioeconomics tells us that the public's sentiment expression affects individual decision-making and hence the market collective decision-making. In this paper, we investigate whether public sentiment drives stock market performance. To be specific, we look at whether there is an association between changes in the Dow Jones Industrial Average (DJIA) and sentiment expression by using a large-scale comprehensive dataset of emotional state swings obtained from Twitter. We analyze relevant textual content on daily Twitter feeds using two sentiment quantification tools: FinBert, which is a categorical indicator that captures positive, neutral, and negative sentiment, and XLNet, which quantifies public sentiment from three types of moods (Positive, Neutral and Negative). Based on the time series dataset of the sentiment indicators, the relationship between public sentiment and DJIA index value is studied through Granger causal analysis and self-organizing fuzzy neural network. In addition, the changes in DJIA closing prices are predicted. Our results show that the accuracy of DJIA predictions can be significantly improved by including information on public sentiment. We have achieved state-of-the-art accuracy when predicting the daily up and down movement of the Dow Jones Industrial Average closing prices.

**Keywords:** Stock market · Sentiment analysis · Time series prediction · Machine learning

## 1 Introduction

Prediction of stock market performance has always been a hot topic and research direction. The Efficient Market Hypothesis (EMH) states that stock market prices in an efficient market follow a random walk pattern since prices reflect all historical and current information. Stock price changes are due to unforeseen future events [1]. The movement of stock prices largely depends on new information coming to the market, such as news posted on the internet and information reported in the financial press. However, future news is highly unpredictable. Hence, stock prices should follow a random walk movement and should never be predicted.

The concept of an “efficient market” has been empirically proved in several early studies [2–5]. The popularity of EMH reached its peak in the eighties [6]. However, the random walk theory has gradually received numerous critics while studies reveal that markets are inefficient in terms of predictability, raising doubts about the assumptions of an “efficient market”. Among these, numerous papers show that existing market anomalies arise from the irrationality of market participants, and stock prices are to some extent predictable due to patterns [6–12].

In addition, recent studies have shown that economic and corporate outcomes can be predicted by early signals could be extracted from online social media, such as Facebook, Twitter feeds, blogs and forums. Empirical evidences demonstrate that online public sentiment are useful in predicting book sales [13], movie sales [14], box-office revenues [15] and a variety of economic indicators [16]. Several studies supports that public sentiment has predictive power of stock price movement [17–19].

In this paper, we test the hypothesis, based on the premise of behavioral economics, that individuals’ emotions influence their decision-making process, leading to a strong correlation between “public sentiment” and “market sentiment.” We perform sentiment analysis on publicly available Twitter data to validate the association between the two. By adopting a model of self-organizing fuzzy neural network (SOFNN), we predict future stock price movements based on the previous days’ Dow Jones Industrial Average (DJIA) index values and sentiment indicators.

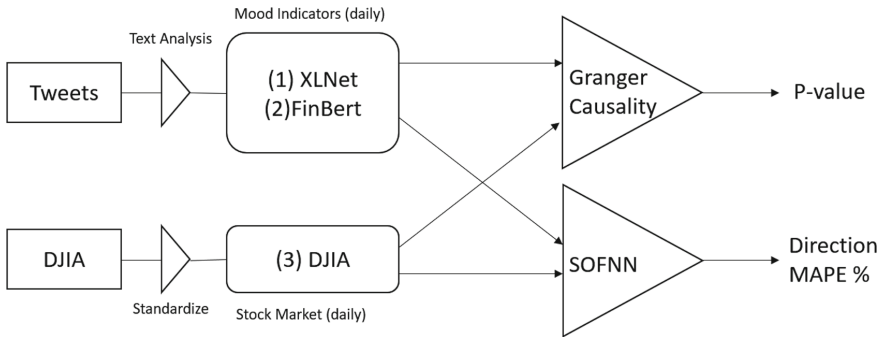
Our work is based on the well-received study by Bollen et al. [19]. The authors predict the closing prices of the DJIA by analyzing the sentiment arising from feeds on Twitter (namely, tweets). The sample dataset of the study includes daily Twitter feeds containing terms that explicitly express users’ mood states. The sample period ranges from February 28th, 2008 to December 19th, 2008. The authors adopt OpinionFinder and Google Profile of Mood States to convert public sentiment into quantifiable values. The resulting time series of mood swings were cross-validated by comparing public sentiment responses to specific cultural events. Then, after verifying the correlation between the sentiment value time series and the DJIA value time series by using Granger causal analysis, the authors used a self-organizing fuzzy neural network, based on sentiment data and historical DJIA data, to predict the direction of changes in the Dow Jones Industrial Index with an accuracy of 86.7%.

Our research combines the experimental results obtained by XLNet and FinBERT to fully exploit the respective advantages of these two algorithms. Get the sentiment label (positive, natural, or negative) of each tweet through each algorithm, and the positive, natural, and negative sentiment values under each algorithm. We use these sentiment labels and sentiment values to predict the up and down trend of the DJIA.

## 2 Related Work

### 2.1 System Design

Please note that the first paragraph of a section or subsection is not indented. The first paragraph that follows a table, figure, equation etc. does not need an indent, either.



**Fig. 1.** Diagram outlining 3 phases of methodology and corresponding data sets.

As shown in Fig. 1, after the dataset is processed, we proceed in three stages. In the first stage, we use 3 sentiment assessment tools on the daily tweet dataset: (1) XLNet, which measures positive, neutral, and negative sentiment from textual content; (2) FinBERT, which measures 3 different sentiments (positive, negative, and neutral) from a textual content dimension; and (3) calculation of the daily sentiment label score. These processes resulted in a total of 12 public sentiment time series, 6 generated by XLNet and 6 generated by FinBERT, each representing a quantified value of public sentiment on a specific date. In addition, we extracted a time series of daily DJIA closing prices from Yahoo Finance. In the second stage, we investigate the hypothesis that public sentiment measured by XLNet and FinBERT can predict future trends in the DJIA. We used Granger causality analysis to correlate DJIA values with the obtained sentiment values. In the third stage, we build a self-organizing fuzzy neural network model to test the hypothesis that the prediction accuracy of the DJIA prediction model can be improved by including public sentiment.

### 2.2 Data Collection

We obtained a dataset of public tweets from January 1st to December 25th, 2010. This data provides the username of the post, the date and time the content was published (GMT+0), and the text content of the tweets (text length is limited

to 140 characters). In the Twitter text dataset, we only consider tweets that contain explicit subjective feelings of their authors' emotional states, such as "I feel", "I am feeling", "I'm feeling", "I don't feel", "I'm", "Im", "I am" and "makes me" [18,20].

### 2.3 Pre-processing

Text data contains more "noisy" words, which do not contribute towards classification [21]. We need to drop those words. In addition, text data may contain tabs, emojis, more white spaces, punctuation characters, stop words, etc [22]. We also need to remove these words. For this purpose, we create our own stop words list, which specifically contains stop words related to finance and general English. After removing stop words, we group all tweets submitted on the same date. To avoid spam, we filter out tweets that contain hyperlinks such as "http:" or "www". In addition, in order to avoid repeated posts affecting the expression of the overall sentiment value, we also remove the tweets with the same content sent by the same users and retain the content and time point of the initial posts. At the same time, we remove the content part of the original tweet in the reposts and retain the text information of the comments left by the users when reposting. Since this study mainly considers the US market, we convert the times of the posts in other time zones to the time zone of the New York Stock Exchange (GMT-8). After processing, the dataset contains 6,809,329 tweets.

### 2.4 Tokenizing Text Mood by XLNet

XLNet uses Transformer XL as a feature extracting architecture, since Transformer XL added recurrence to the Transformer [23,24], which can give the XLNet a deeper understanding of the language context. XLNet is a pretrained model, so we only need to use a fine-tuning method to update the pre-trained model to fit the next task needed.

We randomly select 1000 items from the Twitter data in 2.3 from January 2010 to February 2010 to manually label sentiment labels (Negative, Neutral, Positive). Then we jointly build a training set with the Financial Phrasebank [25] to train the classifier. The Financial Phrasebank is a dataset of sentences from financial news. The dataset consists of 4,840 sentences from English-language financial news categorized by sentiment (Negative, Neutral, Positive) [26,27].

The Twitter Sentiments Dataset [28] is a public dataset. It contains two fields for the tweet and the sentiment label. There are a total of 162,981 sets of data. We randomly select 1,000 of them as the test set to evaluate the performance of the XLNet model. In order to prevent data distortion, the epoch of XLNet is set as 1. The results show that the test accuracy is 0.861, the test loss is 0.23, and the F1-score is 0.87. It meets the needs of our next task.

## 2.5 Sentiment Analysis by FinBert

Although XLNet has excellent features in context understanding and language recognition, more training is required for a larger number of subdivisions in financial-related fields. In order to obtain the accuracy of sentiment value in more subdivided directions, we introduce FinBert [29]. FinBERT is a pre-trained NLP model to analyze the sentiment of financial text. It is built by further training the BERT [30] language model in the finance domain, using a large financial corpus and thereby fine-tuning it for financial sentiment classification. FinBert [31] uses data from Financial Web (6.38B words), Yahoo Finance (4.71B words), and Reddit Finance QA (1.62B words) for pre-training, and related research shows that its text analysis in the financial segment is more accurate. FinBert quantifies the sentiment of tweets in terms of positive, negative, and neutral.

## 2.6 Comparing Sentiment Analysis Results of XLNet and FinBert

To enable the comparison of XLNet and FinBert time series, we standardized them to z-scores on the basis of a local mean and standard deviation within a sliding window of  $k$  days before and after the particular date [32,33]. The principle and mechanism are the same as Gallup’s Economic Confidence Index. The z-score of time series  $X_t$ , denoted  $Z_{x_t}$ , is defined as:

where  $\bar{x}(x_{t\pm k})$  and  $\sigma(x_{t\pm k})$  represent the mean and standard deviation of the time series within the period  $[t-k, t+k]$ . This standardization ensures all time series’ factors to fluctuate around a zero mean and be expressed on a scale of unit standard deviation.

$$z_{x_t} = \frac{x_t - \bar{x}(x_{t\pm k})}{\sigma(x_{t\pm k})} \quad (1)$$

## 2.7 Cross-Validation of XLNet and FinBert Time Series for High-Impact Sociocultural Events

We first validate the ability of XLNet and FinBert to capture various aspects of public sentiment. For this we will apply tweets published during the March period from October 5th to December 5th, 2010. This interval was chosen because it may contain public sentiment reflected by cultural events with significant or complex social impact, namely the US Presidential Midterm Election (November 2, 2010) and Thanksgiving (November 27, 2011). Therefore, the emotion quantification results of XLNet and FinBert can be cross-validated according to the expected responses to these specific events. The time series of emotion values obtained are shown in Fig. 2 and Fig. 3, and expressed as z-score. The formula is shown in Eq. 2.

where  $X$  represents the emotional time series obtained from the 4 groups, which are the sentiment label score of XLNet, the sentiment value of XLNet, the sentiment label score of FinBert, and the sentiment value of FinBert.

$$Y_{Djia} = a + \sum_{i=1}^n \beta_i X_{i-t} \quad (2)$$

From Fig. 2 and Fig. 3, we can see that the sentiment values of XLNet and FinBERT can both respond to the major social events introduced in the study by Bollen et al. [18] and respond to public sentiment.



**Fig. 2.** The XLNet model shows public sentiment swings from tweets posted from October 2010 to December 2010, which can reveal public responses to the presidential midterm elections and Thanksgiving.

**Table 1.** The SSR for each emotion dimension combination is in this table.

Emotion Dimensions	SSR
XLN_score - XLN_value	344.6979244
XLN_score - FinB_value	334.8667083
FinB_score - FinB_value	338.9763747
FinB_score - XLN_value	335.5207981
All	337.9401028



**Fig. 3.** The FinBERT model shows public sentiment swings from tweets posted from October 2010 to December 2010, which can reveal public responses to the presidential midterm elections and Thanksgiving.

The multiple regression results are shown in Table 1. From this table, we conclude that the emotional performance of some FinBERTs is not all consistent with the emotional changes provided by XLNet. The expression of events by the sentiment analysis of a single algorithm cannot well reflect the correlation between public sentiment and special events. If taking all dimensions of emotional changes into account does not give the optimal result, interleaving various dimensions would achieve relatively better results.

## 2.8 The Lag of Public Sentiment on Events

Changes in sentiment values are continuous over time. However, the DJIA series is discontinuous because of the presence of a market closure. We consider the impact of public sentiment on economic changes to be continuous during the market closure. In other words, when the market is closed, the DJIA index just does not show up in the form of data, but the impact of public sentiment is still there. This part of the impact of public sentiment accumulates until the market opens. Alternatively, the DJIA value on the first day after the market opens is not just influenced by one day of public sentiment, but a cumulated expression of public sentiment over several days. Therefore, the average change of the DJIA value from the day before the market closure to the first day of the

market opening is calculated. This average change is used to compute the DJIA value on market closure days. At the same time, a dummy variable is added, with the date of having the actual DJIA recorded as 0 and the date of using the calculated DJIA recorded as 1.

We apply the econometric technique of Granger causality analysis to make a preliminary test on the correlation between DJIA index movement and the daily time series produced by XLNet and FinBert. Granger causality analysis rests on the assumption that the past value of one time series influencing the present and future value of another time series [34]. Granger [35] proposed that the variance of the optimal prediction error of time series X is reduced by including the historical data of time series Y. In fact, this notion is based mainly on predictability but not causality of Y on X [36]. Following Hiemstra and Jones [34], we use linear Granger causality test on the dynamic relationship between daily Twitter sentiment and DJIA index movement.

We thus expect that the lagged values of X exhibit a statistically significant correlation with Y. Correlation however does not prove causation [18]. We are not testing actual causation but whether one time series has predictive information about the other or not. Our DJIA time series, denoted  $D_t$ , is defined to reflect daily changes in stock market value, i.e. its values are the delta between day t and day t-1:  $D_t = DJIA_t - DJIA_{t-1}$ . To test whether our sentiment time series predicts changes in stock market values, we compare the variance explained by two linear models as shown in Eq. 3 and Eq. 4. The first model (L1) uses only n lagged values of  $D_t$ , i.e.  $(D_{t-1}, \dots, D_{t-n})$  for prediction, while the second model L2 uses the n lagged values of both  $D_t$  and the XLNet with the FinBert sentiment time series denoted as  $X_{t-1}, \dots, X_{t-n}$ . Based on Bollen et al. [18], we add the second lag to the sixth lag of  $D_t$  and  $X_t$  in our model L1 and L2.

$$L_1 : D_t = \alpha + \sum_{i=1}^n \beta_i D_{t-i-t} \tag{3}$$

$$L_2 : D_t = \alpha + \sum_{i=1}^n \beta_i D_{t-i} + \sum_{i=1}^n \gamma_i x_{t-i+t} \tag{4}$$

It can be seen from the results of the Granger causality analysis (Table 2), there is a strong correlation between the time series of emotional values and DJIA values. Among them, when t=3, the correlation between sentiment series and the DJIA value series is the highest. In order to show the viewing results more intuitively, we visualize the time series of emotions and the time series of DJIA at t=3. To maintain the same scale, we convert the DJIA delta values  $D_t$  and sentiment value  $X_t$  to z-scores as shown in Eq. 1. And, since the verification shows that the result is better when t=3, we use the data with a lag of 3 days in the model in the subsequent prediction.



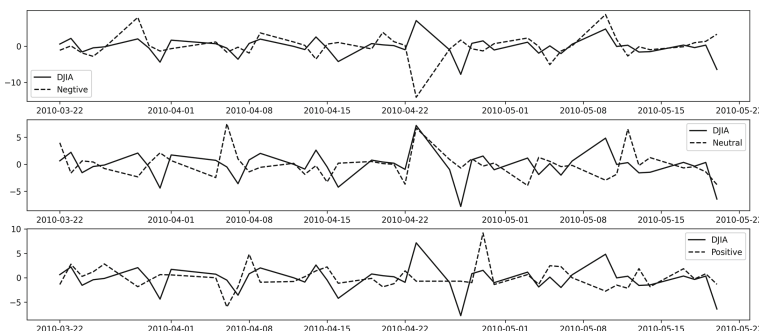
**Table 2.** The p-values of each sentiment value.

Lag	<i>XLNet_value</i>			<i>FinBert_value</i>		
	Negative	Neutral	Positive	Negative	Neutral	Positive
1 day	0.2087	0.0476	0.3504	0.2897	0.3279	0.4833
2 day	0.1276	0.0005	0.0986	0.3626	0.1341	0.3373
3 day	0.0092	0.0025	0.0354	0.0062	0.1088	0.1101
4 day	0.0150	0.0789	0.0584	0.1690	0.0305	0.1247
5 day	0.1479	0.1835	0.2543	0.1041	0.1258	0.0757
6 day	0.3533	0.2959	0.4661	0.3409	0.0178	0.4400

### 2.9 Model Training and Prediction

Since the correlation between sentiment value and DJIA closing prices is non-linear [18], after determining the correlation between lags of Twitter sentiment, lags of DJIA index value and the present DJIA index value, we established a SOFNN model based on the sentiment value and the closing price of the day with a lag of 3 days and 4 days, respectively. We have taken January 8th, 2010 to November 30th, 2010 as the training set, and December 1st, 2010 to December 17th, 2010 as the test set.

The Self Organizing Fuzzy Neural Network (SOFNN) [37] is a 5-layer fuzzy neural network which uses ellipsoidal basis function (EBF) neurons consisting of a center vector and a width vector. Based on the relevant literature, we establish the SOFNN algorithm model. Neural networks have been considered to be a very effective learning algorithm for decoding nonlinear time series data, given that financial markets often follow nonlinear trends [18,38] (Fig. 4).



**Fig. 4.** A panel consisting of three charts. The graph above shows the daily difference in DJIA values (blue: ZDt) versus XLNet’s sentiment values, i.e. negative, neutral, positive, with a lag of 3 days. (Color figure online)

We constructed an online algorithm for SOFNN following the method of paper [39], where neurons are added or pruned from the existing network when new samples arrive. In order to compare the effects of different algorithms on the prediction of the direction of change of DJIA. In contrast to SOFNN, we used logistic regression and SVM. In order to find higher prediction accuracy, we studied 7 permutations and combinations of the input variables of the models, as shown in Eq 5. We finally obtain the prediction results as shown in Table 3.

$$I_{A,B...} = DJIA_{t-k,k-1,k-2,...,1}, X_{A_{t-k,k-1,k-2,...,1}}, X_{B_{t-k,k-1,k-2,...,1}} \dots$$

DJIA t-k,k-1,k-2,...,1 represents the DJIA values and its lagged values. XA,t-k,k-1,k-2,...,1 represents the values of the sentiment dimension and its lagged values. k represents the values of lag days. A, B, C, D represent the dimension of sentiments: the sentiment label score of XLNet, the sentiment value of XLNet, the sentiment label score of FinBert, and the sentiment value of FinBert. I represent the input dataset [40].

Although we can see from Fig. 2 that the changes of each individual dimension of sentiment deviates from the changes of DJIA index values, from the results shown in Table 3, each dimension of sentiment to some degree has contributed on the predictability of the closing values of DJIA. When all sentiment indicators are included, the prediction accuracy reaches the highest, 88.30%. We compute the MAPE value to further test on the accuracy [18], and the results show that the MAPE value is significantly improved.

**Table 3.** The model predicts the upward or downward change direction of the closing price of DJIA compared with the previous day, and compares it with the actual change direction to obtain the accuracy rate.

I	Logistic regression dirction(%)	SVM dirction(%)	SOFNN	
			Dirction(%)	MAPE(%)
-				
<i>XLNet_score-XLNet_value</i>	62.40	67.97	76.21	1.91
<i>XLNet_score-FinB_value</i>	62.40	73.26	88.30	1.55
<i>FinB_score-FinB_value</i>	62.40	64.35	75.86	2.06
<i>FinB_score-XLNet_value</i>	62.40	68.60	85.89	1.73
All	62.40	70.47	80.65	1.59

### 3 Conclusions and Future Work

In this paper, we verify the relationship between public sentiment and DJIA values by surveying a large number of tweets on Twitter. Our results show that, first, public sentiment can indeed be obtained from large-scale tracking through natural language processing techniques in specific situations. Second, the correlation between changes in public sentiment and changes in DJIA values after 3

days was obtained through Granger causality analysis. Third, it is more helpful to improve the prediction accuracy of the DJIA's closing price by the comprehensive inclusion of various sentiment values, rather than just looking at a single dimension of sentiment. Fourth, it verifies the feasibility of XLNet and FinBert in dealing with the influence of text sentiment on market public opinion.

Finally, it is worth mentioning that there are many factors that our analysis did not take into account. First, we observed and screened datasets in specific regions and periods. With the progress of the times and changes in people's lifestyles, further research and verification are needed on the changed Twitter user population and expressions. Second, although we get the results of evaluating public sentiment after validation, there is no objective fact that it can directly reflect public sentiment. That is, we only proved the correlation between emotional state and the prediction result of DJIA value, and there is no data information on the causal mechanism between these two. Third, we currently only consider the one-way effect of public sentiment on changes in DJIA values to make predictions. And the market is complex, and its impact is not just one-way.

About the future work, due to the strong randomness in the expression of public sentiment, more targeted sentiment expression can better reflect the volatility of the stock market. Moreover, there is a certain time lag between public sentiment and stock price volatility, and our results show that an average 3-day lag can best reflect the impact of public sentiment on the market. But this is not the optimal lag period. We find that when public sentiment is more volatile, it takes less time to affect stock prices. Further adjustments to the forecasting model may improve the forecasting accuracy for a wider range of time periods. Therefore, the impact of changes in public sentiment on the market, as well as on investment decisions, remains an area of future research.

## References

1. Fama, E.F.: Efficient capital markets: a review of theory and empirical. *J. Finan.* **25**(2), 383–417 (1970)
2. Kendall, M.G., Hill, A.B.: The analysis of economic time-series-part I: prices. *J. Roy. Stat. Soc.* **116**(1), 11–34 (1953)
3. Larson, A.: Measurement of a random process in futures prices. *Proc. Annual Meet.* **33**, 101–112 (1960)
4. Fama, E.F.: Efficient capital markets: II. *J. Finan.* **46**(5), 1575 (1991). <https://doi.org/10.2307/2328565>
5. Sewell, M. (2022) History of the Efficient Market Hypothesis. Research Note, 11 (04): 04, UCL Computer Science. <https://www.ucl.ac.uk/computer-science/>. Accessed 30 Oct 2022
6. Shiller, R.J.: From efficient markets theory to behavioral finance. *J. Econ. Perspect.* **17**(1), 83–104 (2003)
7. Butler, K.C., Malaikah, S.J.: Efficiency and inefficiency in thinly traded stock markets: Kuwait and Saudi Arabia. *J. Bank. Finan.* **16**(1), 197–210 (1992). [https://doi.org/10.1016/0378-4266\(92\)90085-E](https://doi.org/10.1016/0378-4266(92)90085-E)

8. Kavussanos, M.G., Dockery, E.: A multivariate test for stock market efficiency: the case of ASE. *Appl. Finan. Econ.* **11**(5), 573–579 (2001). <https://doi.org/10.1080/09603100010013006>
9. Schwert, G.W.: Anomalies and market efficiency. *Handb. Econ. Finan.* **1**, 939–974 (2003). <http://www.nber.org/papers/w9277>. Accessed 30 Oct 2022
10. Bouman, S., Jacobsen, B.: The halloween indicator, “Sell in May and Go Away”: another puzzle. *Am. Econ. Rev.* **92**(5), 1618–1635 (2002)
11. Bailey, W., Kumar, A., Ng, D.: Behavioral biases of mutual fund investors. *J. Finan. Econ.* **102**(1), 1–27 (2011). <https://doi.org/10.1016/j.jfineco.2011.05.002>
12. Bartram, S.M., Grinblatt, M.: Global market inefficiencies. *J. Finan. Econ.* **139**(1), 234–259 (2021). <https://doi.org/10.1016/j.jfineco.2020.07.011>
13. Gruhl, D., Guha, R., Kumar, R., et al.: The predictive power of online chatter. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 78–87 (2005). <https://doi.org/10.1145/1081870.1081883>
14. Mishne, G., Glance, N.: Predicting movie sales from blogger sentiment. In: *AAAI Spring Symposium - Technical Report*, SS-06-03, pp. 155–158 (2006)
15. Asur, S., Huberman, B.A.: Predicting the future with social media. In: *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, vol. 1, pp. 492–499 (2010). <https://doi.org/10.1109/WI-IAT.2010.63>
16. Choi, H., Varian, H.: Predicting the present with google trends. *Econ. Rec.* **88**(SUPPL.1), 2–9 (2012). <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
17. Gilbert, E., Karahalios, K.: Widespread worry and the stock market. In: *ICWSM 2010 - Proceedings of the 4th International AAAI Conference on Weblogs and Social Media*, pp. 58–65 (2010)
18. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. *J. Comput. Sci.* **2**(1), 1–8 (2011). <https://doi.org/10.1016/j.jocs.2010.12.007>
19. Goel, A., Mittal, A.: Stock prediction using twitter sentiment analysis. *Cs229.Stanford.Edu*, (December), pp. 1–5 (2012). <http://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>. Accessed 30 Oct 2022
20. Gokulakrishnan, B., Priyanthan, P., Ragavan, T., Prasath, N., Perera, A.: Opinion mining and sentiment analysis on a Twitter data stream. In: *International Conference on Advances in ICT for Emerging Regions (ICTer2012)*, pp. 182–188 (2012). <https://doi.org/10.1109/ICTer.2012.6423033>
21. Choy, M.: Effective listings of function stop words for Twitter. *Int. J. Adv. Comput. Sci. Appl.* **3**(6) (2012). <https://doi.org/10.14569/ijacsa.2012.030602>
22. Kharde, V., Sonawane, S.S.: Sentiment analysis of twitter data: a survey of techniques. *Int. J. Comput. Appl.* **139**(11), 5–15 (2016). <https://doi.org/10.5120/ijca2016908625>
23. Yang, Z., et al.: XLNet: generalized autoregressive pretraining for Language Understanding (2020). [arXiv.org](https://arxiv.org/abs/1906.08237?source=techstories.org). <https://arxiv.org/abs/1906.08237?source=techstories.org>. Accessed 30 Oct 2022
24. Mustapha, M., Krasnashchok, K., Al Bassit, A., Skhiri, S.: Privacy policy classification with XLNet (Short Paper). In: *Garcia-Alfaro, J., Navarro-Arribas, G., Herrera-Joancomarti, J. (eds.) DPM/CBT -2020. LNCS*, vol. 12484, pp. 250–257. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-66172-4\\_16](https://doi.org/10.1007/978-3-030-66172-4_16)
25. Malo, P., et al.: Good debt or bad debt: detecting semantic orientations in economic texts. *J. Assoc. Inf. Sci. Technol.* **65**(4), 782–796 (2013). <https://doi.org/10.1002/asi.23062>

26. Babu, N.V., Rawther, F.A.: Multiclass sentiment analysis in text and emoticons of twitter data: a review. *Trans. Comput. Sci. Comput. Intell.*, 61–68 (2021). [https://doi.org/10.1007/978-3-030-49500-8\\_6](https://doi.org/10.1007/978-3-030-49500-8_6)
27. Adoma, A.F., Henry, N.-M., Chen, W.: Comparative analyses of Bert, Roberta, Distilbert, and xlnet for text-based emotion recognition. In: 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP) (2020). <https://doi.org/10.1109/iccwamtip51612.2020.9317379>
28. Hussein, S.: Twitter Sentiments Dataset. Mendeley Data, V. 1 (2021). <https://doi.org/10.17632/z9zw7nt5h2.1>
29. Colacicchi, L.: Comparison and fine-tuning of methods for financial sentiment analysis (2022). <https://dke.maastrichtuniversity.nl/jan.niehues/wp-content/uploads/2022/01/Colacicchi-Thesis.pdf>. Accessed 30 Oct 2022
30. Gao, Z., Feng, A., Song, X., Wu, X.: Target-dependent sentiment classification with BERT. *IEEE Access* **7**, 154290–154299 (2019). <https://doi.org/10.1109/ACCESS.2019.2946594>
31. Liu, Z., et al.: Finbert: A pre-trained financial language representation model for financial text mining. In: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence [Preprint] (2020). <https://doi.org/10.24963/ijcai.2020/622>
32. Gadri, S., et al.: Sentiment analysis: developing an efficient model based on machine learning and deep learning approaches. In: *Intelligent Computing & Optimization*, pp. 237–247 (2022). [https://doi.org/10.1007/978-3-030-93247-3\\_24](https://doi.org/10.1007/978-3-030-93247-3_24)
33. Dusane, P., Sujatha, G.: Events of interest extraction from forensic timeline using Natural Language Processing (NLP). In: Proceedings of International Conference on Deep Learning, Computing and Intelligence, pp. 83–94 (2022). [https://doi.org/10.1007/978-981-16-5652-1\\_7](https://doi.org/10.1007/978-981-16-5652-1_7)
34. Hiemstra, C., Jones, J.D.: Testing for linear and nonlinear granger causality in the stock price-volume. *J. Finan.* **49**(5), 1639–1664 (1994)
35. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–38 (1969)
36. Shojaie, A., Fox, E.B.: Granger causality: a review and recent advances. *Ann. Rev. Stat. Appl.* **9**, 289–319 (2022). <https://doi.org/10.1146/annurev-statistics-040120-010930>
37. Kuremoto, T., Obayashi, M., Kobayashi, K.: Forecasting time series by SOFNN with reinforcement learning (2015). <https://www.semanticscholar.org/paper/Forecasting-Time-Series-by-SOFNN-with-Reinforcement-Kuremoto-Obayashi/8a5ce65e52077303b8dcbe39a3953219e910ca3f/figure/12>. Accessed 30 Oct 2022
38. Leng, G., Ray, A., McGinnity, T.M., Coleman, S., Maguire, L., Vance, P.: An Interval Type-2 Fuzzy Neural Network for Cognitive Decisions (2014). [https://www.researchgate.net/publication/266849980\\_An\\_Interval\\_Type-2\\_Fuzzy\\_Neural\\_Network\\_for\\_Cognitive\\_Decisions](https://www.researchgate.net/publication/266849980_An_Interval_Type-2_Fuzzy_Neural_Network_for_Cognitive_Decisions). Accessed 30 Oct 2022
39. Leng, G., Prasad, G., McGinnity, T.M.: An on-line algorithm for creating self-organizing fuzzy neural networks. *Neural Netw.* **17**(10), 1477–1493 (2004). <https://doi.org/10.1016/j.neunet.2004.07.009>
40. Nofer, M., Hinz, O.: Using twitter to predict the stock market. *Bus. Inf. Syst. Eng.* **57**(4), 229–242 (2015). <https://doi.org/10.1007/s12599-015-0390-4>