Chapter 8
An Investigation into Influences of Tweet Sentiments on Stock Market Movements

Zhicheng Hao and Yun-Heh Jessica Chen-Burger

Abstract The increasingly common practice by the public of using social media as an information source and basis for decision making for investment has given social media, such as Twitter, a growing influential role in people's behaviour. We investigate the impact of Tweets on the US stock market using sentiment analysis. We examine how the sentiments in Tweets influence the price movements in S&P 500, an excellent indicator for the US stock market and economy in general. We compared the top five influential Twitter outlets and found WSJ, Bloomberg, Forbes, and Reuters consistently showed similar sentiments. We also discovered a significant two trading days' delay in impacting S&P 500 trends where obvious agreements are found between the above outlets. An enhanced version of SentiStrength has been used where we injected finance-related lexicons underpinned by our investor's ontology to obtain a significantly improved and consistent performance.

8.1 Introduction

Social media such as Twitter provides valuable information that reflects public and analysts' opinion and moods and are increasingly used by investors as a tool for making investment decisions [1]. Tweets provide useful insights for smarter decisions on whether one should sell, buy, or hold their shares. However, the stock market is highly volatile and complicated, and stock prices can be largely driven by real-time information that makes them extremely difficult to predict [2]. Twitter is one of the most popular platforms that provide real-time market information and can be automatically extracted to provide features for analysing and exploring hidden financial emotions. Such emotions may indicate market moods leading to stock price changes and volatilities, change volumes of trades, and even cause risks in businesses [3]. However, it is very challenging to accurately understand sentiments from Tweets that

Z. Hao (✉) · Y.-H. J. Chen-Burger
School of Mathematical & Computer Sciences, Heriot-Watt University, Edinburgh, UK
e-mail: zh3@hw.ac.uk

G. Jezic (eds.), *Agents and Multi-Agent Systems: Technologies and Applications 2022*, Smart Innovation, Systems and Technologies 306,
https://doi.org/10.1007/978-981-19-3359-2_8

may include implicit and finance-specific information, e.g., company or macroeconomic information, irregular large volume of transactions, or changes in investors' minds.

In this paper, we describe our efforts to investigate the relationship between the S&P 500 stock price movements and the sentiments as shown in Tweets. S&P 500 is commonly regarded as the best single indicator for the health of the US stock market, as it covers approximately 80% of US market capitalisation. The S&P 500's indexes (SPX) were extracted from Yahoo's Finance portal (yfinance) and the top five financially influential Twitter accounts (WSJ, Bloomberg, Reuters, Forbes, and Donald Trump) were selected for sentiment analysis [4]. The sentiment analysis is implemented by using the lexicon-based sentiment analysis tool SentiStrength [5].

We have chosen to work on data from April 2021 when many significant financial events were discussed on Twitter, e.g., SPX dropped 7% and triggered a trading curb, UK's FTSE 100 fell 10.87%, Frances's CAC 12.28%, and Canadian S&P/TSX 12%. Our research explores the effects of financial related Tweets at the time, then we studied the correlations between these Tweets and SPX.

To boost the capabilities of the general SA analyser SentiStrength, we have developed an Investor's Ontology to capture financially related information, including macroeconomy. We then developed a set of finance-specialism lexicon based on the ontology. This approach enables SentiStrength to show great improvements and consistent performance in analysing finance-related Tweets from an investor's point of view when comparing with the ground truth.

We then deployed the enhanced SentiStrength to analyse Tweets and compare them with movements of SPX. We found very interesting results. We found major financial Twitter outlets reach agreements on their positive or negative sentiments. We also found Tweets of Donald Trump sometimes showed opposite sentiments, but this may be explained by his personal political motivation. Moreover, we detected a consistent two trading days' delay of stock market trend impact, following consensus in Tweets sentiments. It shows media's mood can reverse an upward trend to go down or vice versa. If so, this can be very useful when deciding how to trade in the stock market.

8.2 Background Information

8.2.1 Financial Market

Efficient Market Hypothesis (EMH) [6] declares that prices reflect all information. Investors cannot trade overvalued or undervalued stocks and select a specific timing to sell or buy stocks. However, many researchers and economists found that stock prices can be predicted to some degree [6]. Xu and Berkely [7] evaluated news by using textual information from social media (e.g., Twitter) and found significant correlations between news articles and stock price movements. There are two types of analysis that help investors to outperform the market. First is a technical analysis which explores trading opportunities and predicts future movements by using historical data, such as past stock trends, trading values, and price patterns. The second method is the fundamental analysis based on economic and financial factors, e.g., macroeconomic indexes (e.g., GDP) and microeconomic information (e.g., company performance).

8.2.2 Sentiment Analysis in Tweets and Stock Market

The stock market is highly dynamic and volatile, but it plays a key role in shaping the global economy. Due to its timely publication, social media such as Twitter has become one of the most important sources of information for decision making whether for investment or speculation on the stock market [3, 8]. Chen [1] discovered that articles and commentaries on social media improve the price discovery process and investors' decision-making. Ruiz [9] found important correlation between stock market movements and social media activities. Bollen [3] explored how public mood can affect investors' decision-making and behaviours.

Popular Twitter accounts can greatly influence the stock market, e.g., newsagents (Reuters), well-known companies (Amazon), and individuals (former US President Donald Trump). Tweets allow users [10] multi-directional interaction with companies and have changed the dynamics and nature of corporate disclosure [11]. Companies and public figures can quickly fill information in a vacuum by using their Tweets before rumours, misinformation, and speculation intensify the crisis. Official information can help investors make smarter decisions. Public sentiments and opinions play an important role in human decision-making [3]. Ranco [12] investigated the relationships between 30 stock companies' Twitter accounts and Dow Jones Industrial Average index and discovered the correlation between Twitter sentiments and abnormal returns.

Sentiment analysis (SA) identifies and categorises the emotion and preference of text. However, sentiment analysing Tweets is more challenging compared with conventional text because of their short length, and because they frequently use informal words and symbols that have rapidly evolved. The lexicon-based SA method uses lexicons that are labelled with sentiment polarities [10] that are created manually and are carefully vetted [10]. Generic lexicons can be taken from SentiWordNet [13], WordNet Affect [14], General Inquirer lexicon [15], or LIWC dictionary but they are not labelled with sentiment values [16]. However, SA with a financial background may improve the accuracy of its sentiment classifications. It is therefore interesting to examine the performance of such a SA tool.

8.3 Methodology

We followed a similar methodology as in our previous work [4]. The first step is to identify highly influential Twitter accounts and interesting time periods where important finance-related events have happened for investigation. Secondly, relevant Tweets and stock price data were collected for this duration. The third step includes pre-processing of collected Tweets and stock price data. This is followed by data cleansing and standardisation. Then the automatically generated sentiments and ground truth values are assigned to each Tweet. Sentiment scores are then normalised for exploring their correlations with stock price movements.

8.3.1 Twitter Data Collection

Tweets were extracted via Twitter's API. For each Tweet, we extracted its main text, author_id, creation_date, Tweet_id, and language_used, for the duration of April 2021. We have selected the top five financially influential Twitter accounts: Reuters (23.8 million followers), Wall Street Journal (19 million), Forbes (17 million), Bloomberg (7.5 million), and Donald Trump (88.7 million—before the account suspension). S&P 500's Index was chosen as the target measure that is a weighted index of the market capitalisation of the 500 largest companies in the USA. It represents the overall trend of the US economy and stock market. Yahoo's finance API (yfinance) provides access to SPX. For each trading day, the opening price, hourly closing price, and closing price of the day have been retrieved as indicators to reflect the stock market movements of the day.

8.3.2 Sentiment Analysis and the Ground Truth

SentiStrength is a lexicon-based sentiment analysis method specialising in extracting sentiments from short text [5]. For each text, it detects the pair of negative and positive sentiment polarities: -1 (not negative) to -5 (extremely negative), and 1 (not positive) to 5 (extremely positive). At the heart of this method is the sentiment lexicon that enables SentiStrength to judge the sentiment values of the processed text. Sentiment lexicon is a list of words or phrases that are assigned with emotive semantic orientations, such as the values above.

We used the original and domain-independent lexicon of SentiStrength as a seed and gave it a financial context by adding financially related words, jargon, idioms, and phrases. We then assigned a sentiment value for each vocabulary and phrase for their potential impact on stock market price movements. These financially important and meaningful terms have been taken from an Investor's Ontology that we built based on vocabularies and concepts as presented in financial articles, dictionaries from industry investor's websites, financial research ontologies, and textbooks.

To judge the performance of the newly enhanced SentiStrength, we have defined a set of principles to generate the ground truth from an investor's point of view:

- The decision should be based solely on the possible influence of the Tweets on the US stock market, and not how an individual investor may decide to trade in the stock market.
- The judgement should be made based on the impact on the overall US stock market, and not for a few specific companies.
- The decision should primarily be focused on shorter term (e.g., within 1–7 days), although impacts of some Tweets may be longlasting.
- Use S&P 500 as an indicator for US stock market trends.
- Dramatic news of very large companies will have a bigger impact (esp. in the short-term) on price movements, e.g., Apple Inc. (AAPL) is weighted 6.2% in SPX, so its breaking news will likely show a visible impact on the stock market—such news often affects its business partners, thereby generating rippling effects that amplify its effects in the stock market in the US or elsewhere.

8.4 Experiments

Based on the stock price data collected via Yahoo Finance, we randomly selected 1000 Tweets from each Bloomberg, WSJ, Reuters, Forbes, and Donald Trump as published in April 2021, resulting in 5000 Tweets in total. We have chosen 1000 Tweets per account, so can get a more balanced view across different accounts—news outlets can have far more Tweets per day. These Tweets are also evenly distributed during the trading hours each day. Each Tweet is assigned with a pair of sentiment scores of negative and positive values (Neg, Pos), using SentiStrength and by ground truth. This pair of values is then converted into a single compound sentiment score, SA: SA = Neg, if $|\text{Neg}| > |\text{Pos}|$; SA = Pos, if $|\text{Neg}| < |\text{Pos}|$; otherwise, SA = 0 if $|\text{Neg}| = |\text{Pos}| = 1$. Our newly devised formula, based on our experiments, is superior to the traditional summation of the pair of SA values, as the values of -1 and $+1$ weaken extreme emotions and the range of resulting compound SA is shrunk to $[-4, 4]$ that removes extreme sentiments of -5 and 5. However, -1 and $+1$ can represent no opinion at all.

We determine the sentiment of the entire day based on all Tweets published that day to generate SA_c using our newly devised formulas 8.1 and 8.2. We first determine the general mood of the day, SA_w, that is a weighted SA score based on all Tweets that have expressed an emotion.

$$\text{SA}_w = \frac{\overline{N}_{pos} * N_1 + \overline{N}_{neg} * N_3}{N - N_2} \tag{8.1}$$

where N_1, N_2, and N_3 are the number of positive, neutral, and negative Tweets published that day. $\overline{N}_{pos}$ and $\overline{N}_{neg}$ are the means of the positive and negative sentimental scores. As a result, SA_c shows the compound sentiment score of the day that is determined by the strength of the weighted SA_w of the day, as described below.

$$\mathrm{SA}_c = \begin{cases} if\,\mathrm{SA}_w > 0, \mathrm{SA}_c = \overline{N}_{pos} \\ if\,\mathrm{SA}_w < 0, \mathrm{SA}_c = \overline{N}_{neg} \\ if\,\mathrm{SA}_w = 0, \mathrm{SA}_c = 0 \end{cases} \tag{8.2}$$

The Pearson correlation coefficient (P) was calculated to determine the correlation between the sentiments of Tweets and the stock market price movements. The range is between -1 and 1, where 0 indicates no correlation, 1 indicates a perfect positive linear correlation, and -1 indicates a perfect negative linear correlation.

$$P = \frac{\sum(x_i - \overline{\mathrm{x}})(y_i - \overline{\mathrm{y}})}{\sqrt{\sum(x_i - \overline{\mathrm{x}})^2 \sum(y_i - \overline{\mathrm{y}})^2}} \tag{8.3}$$

where P is Pearson's correlation coefficient, x_i is $\mathrm{SA_c}$ scores in the dataset, $\overline{\mathrm{x}}$ is the mean of all x_i, y_i is stock prices, and $\overline{\mathrm{y}}$ is the mean of all y_i.

8.5 Results

8.5.1 Tweets Sentiment Statistics Analysis

Since the sentiment scores include a pair of negative [-1, -5] and positive [1, 5] values, Fig. 8.1 shows the distributions of the sentiment scores as produced by the SentiStrength and ground truth. Overall, these two methods produced similar results and distributions. Many Tweets are fact-based and therefore evaluated neutral

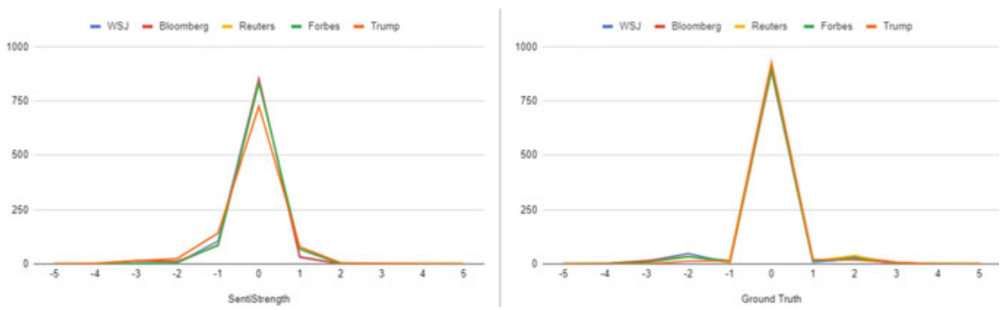

Fig. 8.1 Distributions of compound SA_c scores as generated by SentiStrength (left) and ground truth (right)

sentiments that indicate the attitudes of many (financial) outlets as conservative and do not express positions in their Tweets. However, all Twitter accounts show slightly more negative sentiments. Interestingly, although the account of Donald Trump deploys many positive emotive words, the tool and ground truth, however, have generated more negative sentiments. In addition, many of the "negative" Tweets are judged to be less negative by the tool and the ground truth: the ground truth is also slightly less negative than the tool results. This observation may be explained by the fact that positive emotive words may not necessarily have a positive influence on the stock market if the Tweets include key words/phrases that are perceived to be negative financially. Similarly, negative emotive words may not necessarily generate a negative SA value, if they are not relevant to the financial context, or not having a similar negative impact on the stock market. Overall, Reuters and Forbes exhibit similar sentiments and are more positive than others.

SPX is made available between 9:30 am and 3:30 pm. However, Tweets' publishing time is not limited by this timeframe. To map the effects of Tweets on stock market prices, we have divided a trading day into the following time windows for sampling purposes: 9:30–10:30, 10:31–11:30, 11:31–12:30, 12:31–13:30, 13:31–14:30, 14:31–15:30 (time of closing price), and 15:31–9:29 (9:29 of the next trading day). As a result, we captured the stock market prices at the beginning of a trading day (opening price), the last price at the end of each time window, including the closing price (last price of the day). We also generated the weighted SA_w for all Tweets falling into each of the observed time windows above.

Since the SA_w values and stock prices have different ranges and are therefore not directly comparable, it is necessary for normalisation to place them on the same scale. The min−max normalisation method has therefore been applied to produce the value of N_{diff} for each stock price and SA_w, where N is the variable of concern, and $N_{\min}$ and $N_{\max}$ are the maximum and minimum values in the same time window. The range of N_{diff} is between 1 and 0. Z-score converts a raw value into a normalised value Z by using mean and standard deviations.

$$N_{\text{Diff}} = \frac{N - N_{\min}}{N_{\max} - N_{\min}} (\text{Min} - \text{Maxnormalisation}) \tag{8.4}$$

$$Z = \frac{\text{Value} - \text{Mean}}{\text{Standarddeviation}} (\text{Mean} - \text{SDnormalisation}) \tag{8.5}$$

Table 8.1 shows Pearson correlation coefficients using the normalised N_{diff} of SA_w scores and SPX indices in two trading days' time, as N_{diff} gives a better differentiation. The range of Pearson correlation coefficient can be interpreted as weak

Table 8.1 Pearson correlation coefficient (P) between SA scores and stock market price (in two trading days' time)

	WSJ	Bloomberg	Reuters	Forbes	Trump
S&P 500	0.26	0.33	0.65	0.56	−0.11

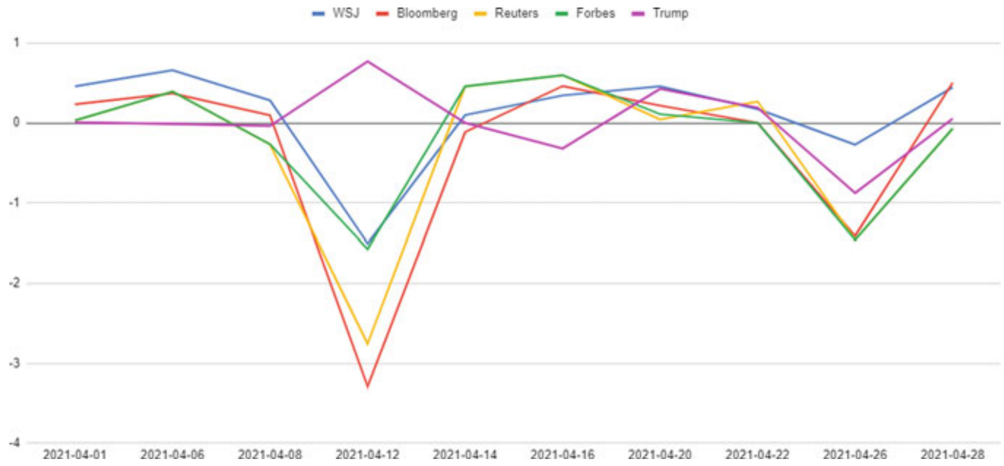

Fig. 8.2 Sentiment distributions of WSJ, Bloomberg, Reuters, Forbes, and Donald Trump

correlation ($0.1 < |P| < 0.3$), moderate correlation ($0.3 < |P| < 0.5$), and strong correlation ($|P| > 0.5$). We can see that all Twitter accounts, except Donald Trump's, have a positive correlation with stock market movements. Reuters and Forbes show relatively strong correlations; Bloomberg shows a moderate one, and WSJ shows a weaker one. Donald Trump's account has a weak and negative correlation with SPX's movements.

In Fig. 8.2, we further compare the normalised sentiments of the top five financially influential accounts. Although Tweets of different accounts contain the opinion of many different investors and analysts, they showed similar opinion trends in the four media outlets (WSJ, Bloomberg, Reuters, and Forbes), except for Donald Trump's account that it sometimes exhibited opposite movements. The strong positive sentiment shown in Donald Trump's Tweets can sometimes link to negative changes in SPX. In fact, when many negative events/news occurred in the stock market, Donald Trump sometimes published opposite positive opinions to encourage investors and stock market. This may explain why the sentiments of Donald Trump's Tweets do not always correspond with that of other news outlets, but may then follow mainstream opinion with one day's delay. In addition, many of Donald Trump's Tweets captured during this time are not directly related to the stock market or finance—as a result, we have decided to exclude Donald's Tweets in our predictive analysis below.

8.5.2 Lagging Effects of Tweets

Increasingly, sentiments of financial related Tweets signal early bullish or bearish trends for stock price movements. It is therefore important to examine these early indicative signals. We analysed the lagging effects of sentiments of Tweets on SPX. According to our analysis, SPX showed a lag of approximately two trading days of impact by the publications of relevant Tweets. In Fig. 8.3, we compared the

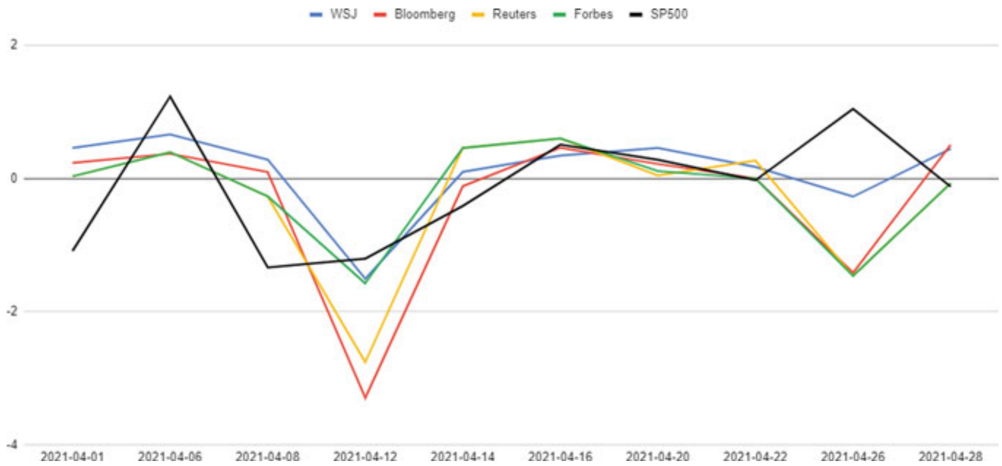

Fig. 8.3 Comparison between sentiments of Twitter accounts and SPX

normalised sentiment values (*Z*-score) of the four financial news outlets (excluding Donald Trump's Tweets) and SPX. *Z*-score is used, as it exhibits a bigger differentiation here. Figure 8.3 shows that a strong dip in commonly shared opinion from news outlets (local minimum) can impact SPX with a (reversed) downward trend in two trading days' time (e.g., 12 and 16 April). Similarly, a commonly shared upbeat opinion (local maximum) can be followed by an upward trend in SPX on two trading days time (e.g., 12 and 16 April). This lagging effect may be explained by the fact that investors need time to digest and process information before making an investment decision. It can also be that information takes time to spread between investors before actions are taken. There are also the "sheep" effects where some investors blindly follow stock market movements.

8.6 Conclusion and Discussion

In this paper, we assessed the impact of financial related Tweets on stock market price movements by using sentiment analysis. We explored the correlation between the opinion of the top five financially influential Twitter accounts and the stock price movements of the S&P 500. All Tweets were collected from Wall Street Journal, Bloomberg, Reuters, Forbes, and the former US President Donald Trump's accounts. The chosen data collection period (1–30 April 2021) is very interesting because many significant events happened during that time: the US–China trade war, crash in stock market (trading curb), and the emergency of Coronavirus. We divided this period into smaller observation time windows (hourly when suitable) to obtain suitable samples for comparison. Within these time windows, we then compared the sentiments as expressed in these Tweets and how they corresponded with SPX's movements.

We use the SA scores as generated by SentiStrength for comparison. The original generic domain-independent sentiment analysis tool, SentiStrength, was enhanced by our set of newly developed financially related lexicon that is underpinned by our Investor's Ontology. This set of lexicons greatly enhances the accuracy of SentiStrength in analysing financially related Tweets. As a result, we found a high level of agreement between the sentiment values as generated by the tool and the ground truth.

In our research, we found the four major news outlets (WSJ, Bloomberg, Reuters, and Forbes) forming similar positive or negative opinions. However, Donald Trump's Tweets sometimes showed opposite opinions. This may be because his Tweets can exhibit political motives that are not consistent with the financial market mood.

We, therefore, compared the sentiments of the above four major news outlets with the movements of SPX. We found a relatively strong and positive correlation between them, especially Bloomberg and Forbes have shown strong correlations. In addition, we found Tweets possessing significant sentiments (i.e., local maximum and minimum) when in agreement with the news outlets, it showed a consistent trend in SPX in two trading days time. However, such impact can sometimes reverse price trends, when the sentiments are extreme, e.g., when many $SA_w > |2|$.

Given the data in the above interesting time period, we found significant correlations between the sentiments in Tweets as published by major news outlets and SPX movements. However, we plan to expand our data set to a smaller granularity or a longer duration to further investigate and determine whether such influences are consistent and whether there is any anomaly. Also, if there is any anomaly, what will they be? Having such further investigation will enhance our understanding of the reliability and predictability of sentiments in Tweets towards stock market movements. We also learned that different news outlets have their own "personalities", i.e., more "conservative" or "positive" than others. However, when we aggregated their opinions to find "significant" opinion spikes, their summated opinions showed consistent predictability towards stock market trends. As not all Donald Trump's Tweets are finance-related, it may therefore be interesting to include other major financial outlets in our research for comparison purposes.

References

1. Chen, H., De, P., Hu, Y.J., Hwang, B.-H.: Wisdom of crowds: the value of stock opinions transmitted through social media. Rev. Financ. Stud. **27**(5), 1367–1403 (2014)
2. Yang, S.Y., Mo, S.Y.K., Zhu, X.: An empirical study of the financial community network on twitter. In: 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), pp. 55–62. IEEE (2014)
3. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. J. Comput. Sci. **2**(1), 1–8 (2011)
4. Hao, Z., Chen-Burger, Y.-H.J.: Analysing tweets sentiments for investment decisions in the stock market. In: Agents and Multi-Agent Systems: Technologies and Applications 2021, pp. 129–141. Springer (2021)
5. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. J. Am. Soc. Inform. Sci. Technol. **63**(1), 163–173 (2012)

6. Malkiel, B.G.: The efficient market hypothesis and its critics. J. Econ. Perspect. **17**(1), 59–82 (2003)
7. Xu, S.Y., Berkely, C.: Stock price forecasting using information from Yahoo finance and Google trend. UC Brekley (2014)
8. Yang, S.Y., Mo, S.Y.K., Liu, A.: Twitter financial community sentiment and its predictive relationship to stock market movement. Quant. Financ. **15**(10), 1637–1656 (2015)
9. Ruiz, E.J., Hristidis, V., Castillo, C., Gionis, A., Jaimes, A.: Correlating financial time series with micro-blogging activity. In: Proceedings of the fifth ACM international conference on Web search and data mining, pp. 513–522. 2012
10. Dhaoui, C., Webster, C.M., Tan, L.P.: Social media sentiment analysis: lexicon versus machine learning. J. Consum. Mark. (2017)
11. Lee, L.F., Hutton, A.P., Shu, S.: The role of social media in the capital market: evidence from consumer product recalls. J. Account. Res. **53**(2), 367–404 (2015)
12. Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., Mozetič, I.: The effects of Twitter sentiment on stock price returns. PLoS ONE **10**(9), e0138441 (2015)
13. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. Lrec **2010**(10), 2200–2204 (2010)
14. Strapparava, C., Valitutti, A.: Wordnet affect: an affective extension of wordnet. In: Lrec, vol. 4, 1083–1086 edn, p. 40. Citeseer (2004)
15. Stone, P.J., Dunphy, D.C., Smith, M.S.: The General Inquirer: A Computer Approach to Content Analysis (1966)
16. Pennebaker, J.W., Mehl, M.R., Niederhoffer, K.G.: Psychological aspects of natural language use: our words, our selves. Annu. Rev. Psychol. **54**(1), 547–577 (2003)