Dancing with Trump in the Stock Market: A Deep Information Echoing Model

KUN YUAN, GUANNAN LIU, and JUNJIE WU, Beihang University
HUI XIONG, Rutgers University

It is always deemed crucial to identify the key factors that could have significant impact on the stock market trend. Recently, an interesting phenomenon has emerged that some of President Trump's posts in Twitter can surge into a dominant role on the stock market for a certain time period, although studies along this line are still in their infancy. Therefore, in this article, we study whether and how this new-rising information can help boost the performance of stock market prediction. Specifically, we have found that the echoing reinforced effect of financial news with Trump's market-related tweets can influence the market movement—that is, some of Trump's tweets directly impact the stock market in a short time, and the impact can be further intensified when it echoes with other financial news reports. Along this line, we propose a deep information echoing model to predict the hourly stock market trend, such as the rise and fall of the Dow Jones Industrial Average. In particular, to model the discovered echoing reinforced impact, we design a novel information echoing module with a gating mechanism in a sequential deep learning framework to capture the fused knowledge from both Trump's tweets and financial news. Extensive experiments have been conducted on the real-world U.S. stock market data to validate the effectiveness of our model and its interpretability in understanding the usability of Trump's posts. Our proposed deep echoing model outperforms other baselines by achieving the best accuracy of 60.42% and obtains remarkable accumulated profits in a trading simulation, which confirms our assumption that Trump's tweets contain indicative information for short-term market trends. Furthermore, we find that Trump's tweets about trade and political events are more likely to be associated with short-term market movement, and it seems interesting that the impact would not degrade as time passes.

CCS Concepts: • **Information systems** → **Deep web**; • **Applied computing** → **Economics**; • **Computing methodologies** → *Information extraction*; Neural networks;

Additional Key Words and Phrases: Stock market prediction, information echoing, Trump, Twitter, deep learning

ACM Reference format:
Kun Yuan, Guannan Liu, Junjie Wu, and Hui Xiong. 2020. Dancing with Trump in the Stock Market: A Deep Information Echoing Model. *ACM Trans. Intell. Syst. Technol.* 11, 5, Article 62 (July 2020), 22 pages.
https://doi.org/10.1145/3403578

This work was supported by the National Key R&D Program of China (2019YFB2101804). G. Liu, J. Wu, and X. Xiong were supported by the National Natural Science Foundation of China (NSFC) (G. Liu: 71701007; J. Wu: 71725002, 71531001, U1636210; H. Xiong: 91746301).
Authors' addresses: K. Yuan and G. Liu (corresponding author), the School of Economics and Management, Beihang University, Beijing 100191, China; emails: {yuankun, liugn}@buaa.edu.cn; J. Wu (corresponding author), the School of Economics and Management, Beihang University, Beijing 100191, China and the Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang, University, Beijing 100191, China; email: wujj@buaa.edu.cn; H. Xiong, the Management Science and Information Systems Department, Rutgers University, Newark, NJ USA; email: hxiong@rutgers.edu.

2157-6904/2020/07-ART62 $15.00
https://doi.org/10.1145/3403578

1 INTRODUCTION

To gain maximum profits and avoid potential losses in the stock market, the stock trend prediction problem has already attracted great attention for many decades in both academia and industry. Although many sophisticated prediction models have been proposed, it is still quite a challenging task due to the highly volatile and non-stationary nature of the stock market. In such an information age, people can easily access information from different sources, resulting in a more volatile market that is sensitive to a variety of information.

Stock market research presents two elemental trading philosophies: technical analysis and fundamental analysis [21, 26]. Technical analysts believe that historic market movements exhibit regularity with modeling some indicators of a stock (e.g., stock price, trading volume), and specific patterns can be distinguished by pattern recognition techniques [19, 24, 31]. However, the performance of these models heavily depends on the selected indicators, which is subject to interpretation of the discovered patterns. In contrast, fundamental analysts attempt to predict the price trend based on economic and business data, with the belief that examining fundamentals such as overall economy, financial conditions, and financial reports provides insights for future market trend prediction [1, 8]. However, Shleifer and Vishny [36] point out that investment decisions rely less on fundamentals than traditional financial theories suggest.

With the development of web media, textual information from breaking news and social media, such as tweets, blogs/microblogs, and discussion boards, enriches the knowledge of investors and also makes them easily influenced by their peers. A large sum of prior work has attempted to extract valuable market information and public moods from different textual media and further utilize them to capture the relationship between high-level market information and stock movements [2, 13, 18, 25]. Preis et al. [29] show that Google search queries could provide early indicators of market movements. Meanwhile, some studies in behavioral finance show that investment decisions are significantly affected by social mood and investor sentiment [11, 27]. Bollen et al. [4] indicate that the specific public mood from Twitter feeds can improve the prediction accuracy of the Dow Jones Industrial Average (DJIA). The success of this stream of studies has revealed that early indicators provided by the social media are correlated with stock market movements and can be further utilized to forecast future economic activity.

The past 2 years have witnessed an interesting phenomenon that Donald J. Trump, the 45th president of the United States, frequently expressed his attitude toward some companies and future policies in finance and trade in Twitter and unexpectedly caused short-term booming or slump of the DJIA. For example, his assault tweets against Amazon on its U.S. Postal Service contract wiped out $5.7 billion of its market value and made the DJIA down as much as 758 points at one point, a brutal first day of trading for the second quarter of 2018 on April 2, 2018. To the best of our knowledge, we are the first to observe the continuous impact from government officials' tweets on the stock market, even on the currency exchange rate [23]. According to the observed frequency and consequences of Trump's tweets, it is reasonable to view Trump's tweets as an auxiliary market-related information source. Owing to the unexpected huge influences, Trump's market-related tweets can provide impactful information, with even greater impact than traditional financial news. According to previous work, the tweets can be regarded as a specific type of online texts that directly influence the market trend. However, different from common online texts that provide indicators of stock movements, Trump's tweets contain more meaningful semantics to explain why and how they can impact the stock market. Moreover, the influences of his tweets may be intensified when they echo with financial news reports. This insight provides the possibility for more accurate stock market movement prediction when capturing predictive market-related features from the echoing effects from both information sources. Therefore, in this article, we mainly concentrate on how to utilize the new-rising information source, along with the traditional financial news, to boost prediction performance with clear interpretation.

Taking a panoramic view of all of Trump's tweets, some tweets on specific topics, such as trade wars and international disputes, can directly move the stock market due to huge influence from the president of the United States. However, this direct impact is slowly reduced since the market gradually gets accustomed to Trump's frequent commenting. In this condition, note that when the impactive tweets and the financial news are discussing the same event, this echoing phenomenon providing information meets or exceeds investors' expectations that can effectively intensify the impact to determine the market trend. In summary, this echoing relationship with financial news dominates the typical way of how Trump's tweets impact the market trend. Furthermore, in many cases, we notice that shares of the companies mentioned by Trump would rebound from their losses on the same day. This observation illustrates that impact from Trump's tweets on the stock market is mostly short lived, and naturally makes it possible for short-term market trend prediction.

Therefore, considering the echoing effects between Trump's market-related tweets and financial news, we need to collectively model the information from both sources with a fusion perspective. Inspired by this, in this work, we design a deep information echoing model through a bi-directional long short-term memory (BiLSTM) network, which models how Trump's market-related tweets echo with financial news to predict the market trend. For the purpose of identifying the most impactful information from the two noisy sources, we design four hierarchical attention strategies to highlight the crucial words in each news title and tweet, and also measure the importance of the news title and tweet to construct efficient representations of each period, respectively. Moreover, we propose a novel information echoing module to learn the high-level market-related echoing representation of the two sources. The designed scheme not only improves the prediction performance but also can explain which and how Trump's tweets impact stock market movements.

In addition, to conduct effective and efficient learning, we propose two learning strategies to guarantee model performance. Confronted with the insufficient training samples problem, we adopt pre-training and fine-tuning techniques to enhance the model training effect. However, noting the fact that the focal topic of Trump's impactive tweets gradually drifts from specific company criticism to trade-related dispute, we introduce a temporal weight function on training samples to fine tune the model parameters during the training process. The ablation experiments demonstrate the effectiveness of the proposed learning strategies.

Furthermore, to validate the effectiveness of the proposed deep information echoing model, extensive experiments have been implemented on the DJIA index based on the information gathered from real-world financial news in *Reuters* and Trump's tweets. Compared with the baseline methods, our deep echoing model performs the best, with accuracy being 60.42%, and also obtains remarkable accumulated profits through a trading simulation. Moreover, we also take a close look at Trump's market-related tweets through the market relevance estimated from the model and find that tweets with topics on finance, trade, and politics are more likely to shake up market movements. These results not only confirm that incorporating Trump's tweets can improve short-term trend prediction accuracy but also validate the effectiveness of the proposed echoing module with some reasonable interpretations.

The remainder of this article is organized as follows. We introduce the related work in Section 2. Section 3 provides insights into how Trump's tweets move the market trend. Then in Section 4, we formally define the market trend prediction problem. Next, Section 5 presents the designed deep information echoing model, including the detailed framework structure and modeling components. Subsequently, in Section 6, we conduct extensive experiments on real-world data. Finally, we conclude the article in Section 7.

2 RELATED WORK

For the sake of gaining maximum profits from the stock market, many efforts have been made toward the stock market trend prediction problem. The data sources used for market prediction cover a wide range of types, including technical indicators, fundamentals, news, and web text from social media.

Technical analysts believe that historic market movements are bound to repeat themselves, so patterns exist that modeling techniques can distinguish [26]. Therefore, modeling techniques such as the ARIMA model [7] and extensions [28], support vector regression [30], neural network [19], and deep learning models [3, 9, 31] have been utilized for financial prediction. However, the efficient market hypothesis (EMH) proposed by Fama [15] states that the stock prices are driven by new information rather than current and past prices. In contrast, fundamental analysts generally depend on fundamentals such as inflation, joblessness, and return on equity to forecast future stock prices [1, 10, 12]. However, Shleifer and Vishney [36] point out that investment decisions rely less on fundamentals than traditional financial theories suggest.

Although market-related news that drives stock movements may be unpredictable, some literature suggests that early indicators extracted from online sources (e.g., news, blogs, and Twitter feeds) can be used to predict movements of various economic indicators [5, 13, 20, 25, 33, 34]. Tetlock [38] found that high media pessimism predicts downward pressure on market prices using the news from the *Wall Street Journal*. Goonatilake and Herath [17] concluded that the number of news articles was related to the market fluctuations of the DJIA and crude oil stock prices. As web information grows, recent work also applied deep learning and NLP techniques to stock prediction [35, 39]. Ding et al. [14] proposed a deep convolutional neural network to model both short-term and long-term influences of events extracted from news titles on stock price movements. Hu et al. [18] designed hybrid attention networks to predict the stock trend with attentively identifying more influential news from the sequence of recent stock-related news.

In addition, studies in behavioral finance also have demonstrated that investment decisions are significantly affected by investor sentiment [22, 27]. Bollen et al. [4] found that public sentiment expressed in large-scale collections of daily Twitter posts are correlated to the changes of the DJIA. Yu et al. [40] demonstrated that sentiments from public news and social media have also been considered to correlate with the stock trend. Si et al. [37] built a semantic stock network from Twitter and proposed a topic-sentiment time-series regression model to predict stock prices.

To summarize the prior work, efficient early indicators provided by the information source play the key roles to achieve accurate stock prediction. More and more web text from social media is demonstrated to drive the stock market just as news. Recently, some researchers have noticed that Trump's tweets on global affairs seem to have a short-term effect on stock price, as now investors always keep an eye on the content that is interpreted as being indicators of future policies. Malaver-Vojvodic [23] studied the impact of Trump's tweets about Mexico and American foreign policy on the daily Mexican peso/U.S. dollar currency exchange rate and found that negative tweets indeed had an impact. Ge et al. [16] analyzed the impact of company-specific tweets of Trump on the stock market and found that the tweets can move company stock prices and increase trading volume and volatility. Born et al. [6] tested the semi-strong form of the EMH by analyzing Trump's tweets about 10 publicly traded firms with standard event study methods. The results show that positive tweets elicit positive abnormal returns on the event date and appear to dissipate over the next few trading days. These studies indicate that Trump's informative tweets are associated with stock movements.

Most prior research merely relies on a single information source to forecast future trends, and thus an integration of news and the new-rising social media content may have the potential to

Table 1. Statistical Results on the Hourly Absolute Variations of the DJIA Index for Two Groups

Statistical Indicators	Tweet Group	Non-tweet Group
Number of samples	1436	2505
Mean of absolute variations	36.7	28.96
Std. of absolute variations	46.57	43.57
Ratio of absolute variations ≥ 100	8.36% (120)	5.55% (139)
Ratio of absolute variations ≥ 200	1.67% (24)	1.2% (30)
Ratio of stock market decline	48.26%	47.09%
Ratio of stock market decline-positive	47.90%	—
Ratio of stock market decline-negative	47.40%	—
Ratio of stock market decline-neutral	49.47%	—
Kolmogorov-Smirnov test	2.5e-9***	

***, p-value < 0.001.

make better prediction. Furthermore, due to the huge influence from the president of the United States, some of Trump's opinions expressed in Twitter would undoubtedly impact investor sentiment. Meanwhile, considering the fact that Trump's tweets can echo with financial news, the two sources of information can collectively provide effective market-related information for prediction. To the best of our knowledge, this work is the first to fuse the two to predict changes in the DJIA.

3 EMPIRICAL ANALYSIS OF TRUMP'S TWEETS

In this section, we first investigate whether the stock market would fluctuate abnormally when Trump posts tweets. Furthermore, we want to reveal the underlying mechanism of how the tweets can impact the market trend. To answer these questions, we carefully examine the relationship between all of the presidential tweets and stock market movement.

3.1 Do Trump's Tweets Shake Up the Stock Market?

To answer the question of whether Trump's tweets indeed shake up the stock market, we classify the data into two groups—a Tweet group and a Non-tweet group—based on the existence of Trump's tweets in that time period and conduct comparative statistical analysis to illustrate the distinction as shown in Table 1. First, we explore the variation of the DJIA before and after Trump posts tweets through calculating the hourly absolute variations of the closing DJIA index. The difference in the calculated mean and standard deviation value indicate that the data in the two groups come from significantly different distribution, which can be further confirmed by a Kolmogorov-Smirnov test with the p-value being 2.5e-9. It reveals the fact that the stock market would fluctuate more fiercely when Trump posts tweets. Moreover, concerning the drastic hourly fluctuation with more than 100 or 200 points, the Tweet group accounts for a higher proportion than the Non-tweet group, which means that Trump's tweets increase the volatility of the stock market.

We also consider whether Trump's tweet bring more positive or negative influence to the direction of variations. Similar to the analysis of absolute variations, we replace it with the direction of variations. In consequence, we find that the ratio of stock market decline increases slightly from 47.09% to 48.26%. Both of preceding two are also in accordance with our observations that the market would rise or fall sharply within a short time due to Trump's tweets in some cases. In addition, we further analyze the correlation of the sentiment of Trump's tweets with the direction of variations. We classify the sentiment of each tweet into positive and negative categories

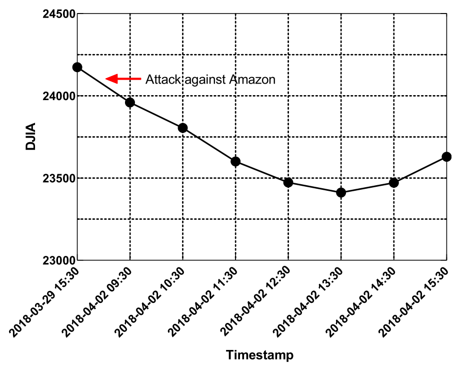

Fig. 1. Attack against Amazon shocks the market.

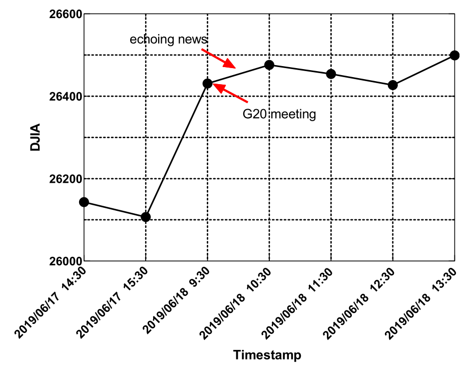

Fig. 2. Good news stimulates the market.

based on the TextBlob library[1] and further conclude the sentiment of one time period into three categories—positive, negative, and neutral—according to the majority rule over all tweets in that same period. As a result, the ratios of the stock market decline are 47.90%, 47.40%, and 49.47% corresponding to three sentiment categories. From the results, we infer that the sentiment cannot provide a discriminative feature to the market trend prediction. In summary, based on the preceding statistical results, we conclude that the short-term fluctuations of the DJIA are correlated with Trump's tweets.

3.2 How Do Trump's Tweets Affect the Market?

Trump is renowned for using Twitter as a strategic tool to express his opinions on global affairs and publicly berate political leaders. His all-embracing tweets cover topics including trade, economics, diplomacy, entertainment, critical comments, and so on. Among them, we find that two typical kinds of tweets—company-specific and trade protectionism—are more prone to shake up the market.

In earlier times, Trump posted tweets to blame some companies, such as Amazon, Google, Facebook, and Boeing, for behaving against his will. Although these comments were oriented to one company specifically, the criticism would sometimes be overinterpreted as a sign to the whole sector and could trigger panic in the market, which finally led to fluctuation of the stock market. Taking Amazon as an example, Trump threatened to accuse the company of not paying its fair share of taxes and rates to the U.S. Postal Service on April 2, 2018, which worried the shareholders and Wall Street investors broadly. Subsequently, Amazon lost more than $50 billion of its market value the next day. Meanwhile, the slump also dragged the corresponding DJIA, dropping nearly 800 points at one point as shown in Figure 1.

Subsequently, Trump began to shift his focus on trade disputes between the United States and China, Europe, and Mexico, among others, and posted tweets about tariffs on steel, European autos, and all kinds of imports from China. In the beginning, there was no doubt that the trade tension would tumble the stock market in fear of trade war. However, the stock market has gradually become accustomed to trade war–related tweets when confronting with longstanding dispute, and as a result such tweets are not guaranteed to dominate the short-term movements compared with other factors as time goes by. Meanwhile, we also notice an interesting phenomenon that some contextual financial news may echo with the market-related information contained in Trump's

[1] https://pypi.org/project/textblob/.

tweets, which provides clear clues to reveal that the information can indeed surprise the market or further intensify market panic, and thus it has more potential to drive the future direction of the market. For instance, as shown in Figure 2, Trump posted a tweet saying "Had a very good telephone conversation with President Xi of China. We will be having an extended meeting next week at the G-20 in Japan" on June 18, 2019, which reset expectations for trade negotiations and pushed the DJIA 352 points up in the morning immediately. Closely followed by that tweet, relevant articles were also widely spread on the Internet, "cheering financial markets on hopes that an escalating trade war between the two countries would abate."[2] In the context of continuous U.S.-China trade tension, investors had been waiting for signs of a possible meeting between the two leaders that could deescalate the recent tension. In this case, the tweet expressed positive information echoing with subsequent media speculation and meeting investors' expectations, which is shown to surge the market.

Therefore, to discriminate whether a tweet has the potential to move the market, we need to consider the intrinsic content of the tweet, and further combine it with news context and analyze the echoing effect. This discovered echoing effect motivates us to design a novel model to combine both information sources from tweets and financial news.

4 PROBLEM STATEMENT

Observing that Trump's tweets have the potential to impact the short-term stock market trend, we choose the DJIA as the prediction target rather than other indexes since the comprised 30 blue-chip American corporations engage in business all over the world and better capture the fluctuations due to changed policy. Hence, in this work, our goal is to conduct a more accurate and fine-grained stock market trend prediction on the hourly fluctuation of the DJIA when Trump posts new tweets via an information echoing perspective.

Assume the sequential hourly financial news of length T as $F_1^T = \{F_1, F_2, \ldots, F_t, \ldots, F_T\}$, where $F_t = \{f_{t1}, f_{t2}, \ldots, f_{tk}, \ldots, f_{tk_t}\}$ denotes a set of news with size k_t posted between time $t-1$ and t. In particular, previous experiments have shown that additional content information cannot contribute more to market prediction than only using the news title; therefore, in this work, we take the news title as raw input with the title f_{tk} described by m_k words—that is, $f_{tk} = \{w^F_{tk,1}, w^F_{tk,2}, \ldots, w^F_{tk,m_k}\}$.

Similarly, we use P_t to denote the j_t tweets posted by Trump between time $t-1$ and t with $P_t = \{p_{t1}, p_{t2}, \ldots, p_{tj}, \ldots, p_{tj_t}\}$, and then we adopt $P_1^T = \{P_1, P_2, \ldots, P_t, \ldots, P_T\}$ to represent the sequential tweet set in the corresponding time period. In addition, each post p_{tj} is assumed to contain n_j words—that is, $p_{tj} = \{w^P_{tj,1}, w^P_{tj,2}, \ldots, w^P_{tj,n_j}\}$. Note that for some time periods, $P_t = \varnothing$ when Trump does not post any tweet.

Finally, we denote $Y_1^T = \{y_1, y_2, \ldots, y_t, \ldots, y_T\}$ to indicate the market trend of the DJIA from time 1 to T, where $y_t \in \{1, 0\}$, and 1 means the stock rises up and 0 means it goes down at time t, depending on the relative difference between the current index and the former one of $t-1$. Similar to many previous stock prediction works, we regard the stock trend prediction problem as a binary classification problem, and our goal is to predict the future trend of the DJIA an hour after we observe the recent presidential tweets—for example, we observe the newest Trump tweets between 9:30 and 10:30 and then predict the future DJIA trend at 11:30 compared with that at 10:30. Then the stock market trend prediction problem can be formulated as discussed next.

[2]https://www.reuters.com/article/us-usa-trade-china/us-china-rekindle-trade-talks-ahead-of-trump-xi-g20-meeting-idUSKCN1TJ1R7.

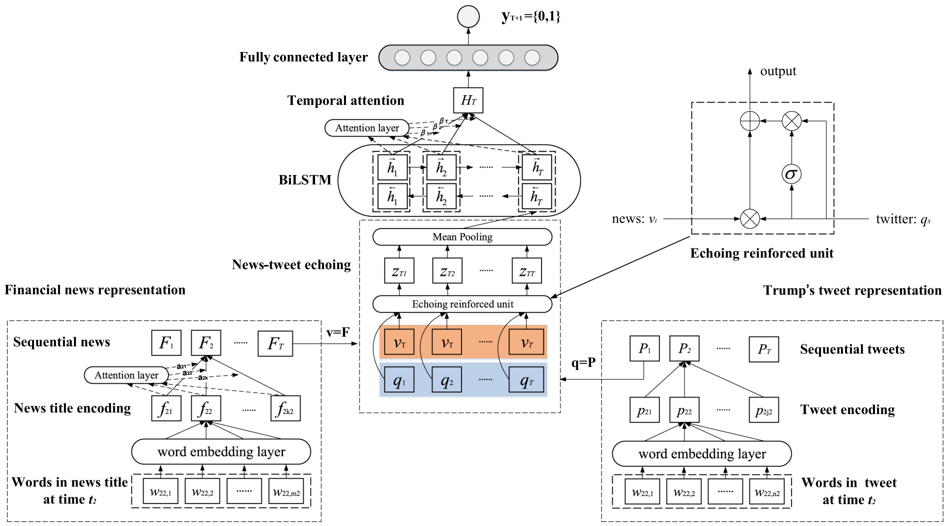

Fig. 3. The framework of the deep echoing model. The echoing reinforced unit in the upper right corner denotes the process of how Trump's tweets echo with news to impact the stock market.

Stock market trend prediction problem. Given the financial news sequential corpus F_1^T with the length T and Trump's posts P_1^T within the same time periods, the goal is to predict the future trend of the DJIA $y_{T+1} \in \{1, 0\}$—that is, up or down an hour after time T subject to the condition $P_T \neq \varnothing$ when we observe Trump's tweets at time T.

5 DEEP INFORMATION ECHOING FRAMEWORK

As analyzed in Section 3, some of Trump's tweets have the potential to impact the short-term movements of stock. Furthermore, when the impactive tweets echo with contemporaneous contextual news, an echoing reinforced effect would move the stock market more extensively. Therefore, in this work, we first extract market-related information from the two raw data sources separately, then the echoing effect is modeled to generate high-level echoing information representations. Finally, the sequential echoing representations are employed to predict the future market trend. The overall network structure of our proposed deep information echoing framework is displayed in Figure 3, which mainly consists of four components—news representation, tweet representation, news-tweet echoing, and sequential modeling—to hierarchically extract high-level market-related information representations from the two data sources for market trend prediction.

Specially, we first pre-train a word representation layer through a word embedding technique on the combined financial news dataset and Trump's tweets. Regarding *news representation*, we adopt word-level and title-level attention to hierarchically extract market-related information through highlighting keywords in each title and further distinguishing driving news titles among chaotic content. Similarly, we also feed the word-level and tweet-level attention into hierarchical *tweet representation* and catch the influential tweet information. Afterward, we design an echoing module to capture the reinforced impact from both sources of information. Then, given all of the temporal sequence of echoing vectors, a BiLSTM network and another temporal attention mechanism can be utilized to summarize the overall market-related echoing information. Finally, the jointly learned representation is fed into a fully connected layer, which maps the input to a binary label to predict the market trend of the next timestep.

5.1 News Representation

Note that even though all of the news comes from a business channel, the quality of each article varies tremendously, and a large portion of online news provides limited information that can barely impact the stock market. Hence, this phenomenon urges the necessity of distinguishing high-quality news as input for further modeling. Generally speaking, a news title contains informative keywords and can better represent the ideas that the news aims to convey. Therefore, to capture the driving information from a series of a chaotic news title corpus in each timestep, two hierarchical attention mechanisms are applied to the news title inputs: word-level attention and title-level attention. Through this hierarchical architecture, the model is expected to encode accurate and informative news vector representations from raw input.

5.1.1 Word-Level Attention. It is intuitive that the meaning of financial news is dominated by several relevant keywords. Therefore, each news title representation can be summarized by a weighted mean of words according to their importance, and correspondingly the word-level attention can serve as the measurement of the importance of each word.

For the kth financial news title f_{tk} with m_k words in news corpus set F_t at time t, we use a pre-trained word embedding layer to transform each word to a real-valued vector—that is, $f_{tk} = \{w^F_{tk,1}, w^F_{tk,2}, \ldots, w^F_{tk,i}, \ldots, w^F_{tk,m_k}\}$. Then we estimate the attention weights with a one-layer network and calculate attention score d with a softmax function as follows:

$$d_{tk,i} = \frac{\exp\left(u_{tk,i}\right)}{\sum_{m=1}^{m_k} \exp\left(u_{tk,m}\right)}, \tag{1}$$

$$u_{tk,i} = {v_d}^T \tanh\left(W_d w^F_{tk,i} + b_d\right), \tag{2}$$

where v_d, W_d, b_d are parameters needed to be learned during the training process. Thus, the attention score $d_{tk,i}$ can be viewed as the importance of the ith word in the title f_{tk}, and the representation vector of the title is encoded by

$$f_{tk} = \sum_{i=1}^{m_k} d_{tk,i} w^F_{tk,i}. \tag{3}$$

5.1.2 Title-Level Attention. For a set of news corpora, the content of news contributes unequally to predict the market trend. For instance, a news report about a specific company cannot be compared with the document of government policy on the market. Therefore, we apply the title-level attention mechanism to summarize all news title vectors within a time period as a single representation identifying the most important news.

The news corpus $F_t = \{f_{t1}, f_{t2}, \ldots, f_{tk}, \ldots, f_{tk_t}\}$ at time t learned in word-level attention can be fed into another one-layer network to learn the importance of each piece of news when used to represent information of the current period. Specifically, we take each news title f_{tk} as input and calculate the corresponding attention score e_{tk} with a softmax function,

$$e_{tk} = \frac{\exp\left(u_{tk}\right)}{\sum_{m=1}^{k_t} \exp\left(u_{tm}\right)}, \tag{4}$$

$$u_{tk} = {v_e}^T \tanh\left(W_e f_{tk} + b_e\right), \tag{5}$$

where v_e, W_e, b_e are parameters learned during the training process. Hence, through the title-level attention mechanism, we are able to distinguish the impactful news, and then the summarized representative vector of current time can be calculated through a weighted means as follows:

$$F_t = \sum_{k=1}^{k_t} e_{tk} f_{tk}. \tag{6}$$

In addition, there other summarization methods exist, like using sum pooling and average pooling, which treat all news equally without discrimination. Although these settings are simple with introducing less parameters, they lead to important and driving information submerged in the chaotic corpus and lose the discrimination ability to market trend.

5.2 Tweet Representation

Similar to news representation, we need to construct tweet representation vectors to summarize relevant market information from Trump's tweets at any time t as well. In view of the difference between tweets posted in the same period (i.e., a tweet about trade war is valued more than that of criticizing other political leaders), a tweet content attention mechanism is used to distinguish valuable information from the various tweets. Therefore, like the modeling process in Section 5.1, two attention mechanisms—word level and tweet level—are applied in tweet representation. Then we only briefly introduce the mathematical formulas for tweet representation since it is essentially the same as news representation.

Assume that tweet set P_t contains j_t tweets posted by Trump from time $t-1$ to t—that is, $P_t = \{p_{t1}, p_{t2}, \ldots, p_{tj}, \ldots, p_{tj_t}\}$. Correspondingly, each post is denoted as $p_{tj} = \{w^P_{tj,1}, w^P_{tj,2}, \ldots, w^P_{tj,n_j}\}$, which consists of n_j words. Denoting L_{word} and L_{tweet} as the word-level and tweet-level attention mechanism network architecture, respectively, the representation vectors of posts at time t can be computed as follows:

$$p_{tj} = L_{word}\left(w^P_{tj,1}, w^P_{tj,2}, \ldots, w^P_{tj,n_j}\right), \tag{7}$$

$$P_t = L_{tweet}\left(p_{t1}, p_{t2}, \ldots, p_{tj_t}\right), \tag{8}$$

where there exist another two groups of parameters v, W and b that need to be learned in the training process. In this way, the distinguished information of tweets in each period is contained in the final representation vector.

5.3 Modeling the Echoing Effect of News with Tweets

After learning the representation vectors of news titles and tweets in each time period, traditional methods would directly concatenate the two vectors and feed them into a fully connected layer for classification, which, however, ignores the interaction between Trump's tweets and financial news and hence the reinforced echoing effect cannot be captured. Therefore, in this work, we propose a novel echoing module to model how Trump's impactful tweets echo with contextual news to drive the short-term stock market.

Moreover, through case analysis, we find that Trump's tweets and news seldom mention one event in exactly the same timestep. In other words, the news may report the event ahead of Trump or later than him. For instance, if Trump posts a tweet at time t, the same event discussed in news can be reported at time $t+1$ or $t-1$. Confronted with such misalignment between the two information sources, we assume that the news F_t has the possibility to report relevant information in Trump's tweets at any timestep from 1 to T. This assumption motivates us to separately model the echoing reinforced effect of each tweet with news F_t between 1 and T, and further aggregate all echoing vectors into a single one as the representations of information in time period t. The detailed two-step modeling process is introduced next.

5.3.1 Echoing Reinforced Impact Unit. As shown in Section 3, when the impactful tweets echo with the contextual financial news, this echoing phenomenon shows a reinforced effect on market-related information and has more potential to move the market. Therefore, we propose an echoing impact unit to capture such echoing effect when given a pair of Trump's tweets and financial news.

For a tweet representation vector P_s in time s, we directly apply a linear transformation to obtain a market-related representation q_s. Here, for simplicity, we let $q_s = P_s$, and we similarly let $v_t = F_t$ for news transformation.

Considering that some tweets posted by Trump already contain sufficient impactful market-related information and they can directly affect the trend of the stock market, we hence propose a gate to control the intensity of the direct impact,

$$p_s = \sigma\left(W_p q_s + b_p\right), \tag{9}$$

where W_p and b_p are parameters of the linear layer. Thus, for any given tweet q_s, p_s represents the proportion of market-related information in the tweet and can also be referred to as the correlation with the market, since when the tweet is more relevant to the market, the larger the p_s is, the more market-related information is retained through $p_s q_s$. This design enhances the modeling interpretation by indicating the extent a tweet correlates with the market.

However, the proposed unit can further control how the tweets can echo with the financial news through an element-wise product $q_s \odot v_t$. Through this mechanism, the echoing effect can be captured when relevant information is discussed in both sources, which would be more important to predict the market trend.

To sum up, the prediction of the stock market trend from Trump's tweets can be viewed as a mixture of market-related information contained in the tweets and the echoing effect arising from the related financial news, which yields the following impact unit:

$$z_{ts} = p_s q_s + q_s \odot v_t, \tag{10}$$

where z_{ts} represents the echoing reinforced representation of the tweet q_s with news v_t.

5.3.2 Echoing Information Aggregation. Then we utilize Equation (10) to summarize the echoing reinforced effect of all Twitter vectors P_1^T on specific news F_t and aggregate those echoing vectors into a single one with an averaging pooling layer as the final representative information at time t—that is,

$$z_t = \sum_{1 \le s \le T} z_{ts}/T. \tag{11}$$

Specifically, when no tweet exists at time t, we retain the news vector; otherwise, the echoing vector is encoded with Equation (10)—that is,

$$z_{ts} = \begin{cases} p_s q_s + q_s \odot v_t & q_s \neq \varnothing \\ v_t & q_s = \varnothing \end{cases} \tag{12}$$

such that the final calculated representative vector contains the market-related content of Trump's tweets and echoing effect of the two sources. With improvement in the interpretation of the proposed model, it further enables us to understand whether and how Trump's tweet cooperates with financial news to determine the market trend.

5.4 Sequential Modeling

Considering that there exists temporal dependence in the sequence of a corpus across all time periods, an investor prefers to judge the future trend based on a unified global context rather than solely relying on any single piece of information. This manner inspires us to consider the dependency of each time period upon all of the others, which could more synthetically assess the comprehensive influence to the market trend. Therefore, to imitate the analysis process of humans, we leverage a BiLSTM network to encode the temporal effect instead of a vanilla LSTM that merely models the forward dependency.

Suppose that we obtain the temporal sequential representative vectors $Z_1^T = \{z_1, z_2, \ldots, z_T\}$, and the hidden encode vector h_t is the concatenation of the forward hidden vector $\overrightarrow{h_t}$ and backward hidden vector $\overleftarrow{h_t}$ at t-step. Thus, we have

$$\overrightarrow{h_t} = LSTM(z_t, \overrightarrow{h_{t-1}}), \tag{13}$$

$$\overleftarrow{h_t} = LSTM(z_t, \overleftarrow{h_{t+1}}), \tag{14}$$

$$h_t = [\overrightarrow{h_t}, \overleftarrow{h_t}]. \tag{15}$$

The encoded h_t incorporates the information of both its surrounding context and itself. Furthermore, since the information at different time periods contributes unequally to the market trend, we adopt a temporal attention mechanism to summarize the discriminative information across all timesteps and generate the overall representative vector H_T, which is calculated as follows:

$$\beta_t = \frac{\exp(u_t)}{\sum_{m=1}^{T} \exp(u_m)}, \tag{16}$$

$$u_t = v_h^T \tanh(W_h h_t + b_h), \tag{17}$$

$$H_T = \sum_{t=1}^{T} \beta_t h_t, \tag{18}$$

where v_h, W_h, and b_h are parameters learned in the training process. Thus, the final H_T is the weighted sum of sequential corpus information according to its importance of contained information.

Trend prediction. With the hierarchical echoing network, we jointly learn the overall representation vector H_T from the two sources. Then we feed it into a fully connected layer to predict the market trend label $y_{T+1} \in \{0, 1\}$ at the next timestep.

5.5 Training Process

Given the proposed prediction framework, there remain some challenges to train the deep echoing model. Since Trump does not post tweets all the time and only 36.43% of data samples contain the recent presidential tweets, the number of the training set $D = \{(F_1^T, P_1^T) | P_T \neq \varnothing\}$ is limited and insufficient to learn a complex deep model. Moreover, we also observe the phenomenon that the focal topic of Trump's tweets impacting the stock market varies with the passing of time, so a well-trained model with former training samples is not guaranteed to perform well in future testing cases. Therefore, to conduct efficient and effective learning, the echoing model must cope with the small sample training problem and meanwhile consider the drifting sample distribution.

Confronted with the problem of insufficient training samples with Trump's tweets, we propose to leverage the pre-training and fine-tuning techniques to train the model, which consists of two supervised training phases. During the pre-training phase, we consider to augment the training set by incorporating other training samples $\{(F_1^T, P_1^T) | P_T = \varnothing\}$ into model training by viewing Trump's tweets as a zero vector. Then, on the basis of the pre-trained model, we fine tune the model with an early stop merely relying on the training set D. Through the experiment, we find that the pre-training technique can effectively enhance model performance by reducing overfitting on a small training set.

In addition, since the focal topic of Trump's tweets gradually changes from company-specific criticism to trade wars against other countries with the passing of time, the model also need to emphasize future samples rather than previous ones to capture this sample distribution drifting problem. Therefore, we apply a weight function $\phi(t)$ to any training sample i so that the echoing model is forced to learn this topic drifting problem by assigning importance weight λ_i according to the posted timestamp t_i of a tweet. For the sake of simplicity, we propose a linear function to approximate the drifting effect as follows:

$$\lambda_i = \phi(t_i) = c + (t_i - t_0) / T_0, \tag{19}$$

where t_0 denotes the Unix timestamp of the starting training time on January 20, 2017, and $t_i - t_0$ represents the elapsed seconds between the starting time and the ith data. c is a constant and $c > 0$, which can be obtained by a grid search strategy. Note that the training data is organized from January 2017 to August 2018, which lasts for 20 months. Therefore, we construct $T_0 = 3{,}600 \times 24 \times 30 \times 20 = 51{,}840{,}000$ (20 months) to normalize the Unix timestamp of each training data into $[0, 1]$. In this manner, the overall weight is transformed into the range of $[c, 1 + c]$. Notice that the smaller c is, the greater weight is assigned to the recent training sample. In this regard, as time passes, the subsequent training sample weighs higher, which captures the drifting trend and thereby can achieve better performance on the testing set.

6 EXPERIMENTS

In this section, we first introduce the experimental settings. Then we conduct experiments on real-world data to evaluate the performance of our proposed deep echoing model. Furthermore, several cases are shown to interpret the modeling ability of our model and reveal the impact of Trump's tweets on the market trend.

6.1 Experimental Setup

6.1.1 Dataset. We collected 74,747 pieces of financial and economic news with titles, content, and timestamps from *Reuters*[3] from October 2015 to April 2019. Most news reports are about market policies and company affairs. As previous research has indicated that news content cannot provide additional information to further improve the prediction accuracy [13, 35], we only regard the news titles as the input of the news corpus in the experiment.

Meanwhile, we collected the closing DJIA index points and stock prices of 30 comprised blue-chip American corporations from January 2017 to April 2019 from *Yahoo Finance*,[4] which were used as input to some baseline methods. The frequency of the data collections is hour by hour from 9:30 to 15:30—that is 9:30, 10:30, ..., 15:30—according to trading time of the stock market. Then, the corresponding closing DJIA index points were transformed to binary values (i.e., rise or fall) as predictors based on the difference of points between the current and next time period.

Moreover, we crawled Trump's tweets from the Trump Twitter Archive[5] from January 20, 2017, to April 18, 2019, since his inauguration as president of the United States. To align the timelines of the three sources, we took the timeline of the DJIA indexes containing 3,942 data samples as standard. In other words, we divided the financial news and tweets into seven segments: 9:30 to 10:30, 10:30 to 11:30, 11:30 to 12:30, 12:30 to 13:30, 13:30 to 14:30, 14:30 to 15:30, and 15:30 to the next 9:30. Along these lines, all financial news and tweets posted out of the trading hours were also collected as a corpus like that of the 1-hour period in trading hours and used as features for subsequent

[3] https://www.reuters.com/finance.
[4] https://finance.yahoo.com.
[5] http://www.trumptwitterarchive.com.

prediction. In the following experiments, we used the data from January 20, 2017, to August 15, 2018, with 2,759 data points as the training time period and the following 1,183 data points from August 16, 2018, to April 18, 2019, as testing time period. Furthermore, we randomly sampled 10% of the training data as the validation set, so as to choose several epochs with the best performance in the pre-training phase. Among the divided sets, the number of time periods containing recent Trump's tweets are 786, 137, and 513 in the training, validation, and testing set, respectively. Since our target is to explore whether the auxiliary Trump tweets can help boost market trend prediction in a short time, we use the 513 testing Twitter samples to evaluate the performance for all prediction models, and the positive samples (rising trend) occupy 52.04% in the testing set. Meanwhile, the models are optimized with the loss function being the binary cross entropy.

We also pre-trained a word-level embedding using the titles and content of collected financial news and some tweets in Twitter as the corpus. We tokenized each news title and tweet, and we removed the stop words and low-frequency words used fewer than five times. Then, a 100-dimension word embedding layer was obtained using the Gensim library [32]. The trained word embedding was applied to encode both news titles and Trump's tweets into dense vector representations as input to all deep learning models.

6.1.2 Baselines. To evaluate the performance of our proposed deep information echoing model, we compare the results with the following methods:

- *Naive Counting*: A naive thought is that the future trend obeys majority rule in several recent time periods. For instance, if three of five recent time periods show a rising trend, then the future trend is rising.
- *GBDT*: Furthermore, we also train a GBDT model by taking several recent market trends as input to examine whether some patterns exist.
- *LSTM-News*: We introduce an LSTM-based model for news-oriented market trend prediction, which is a prototype in the work of Hu et al. [18]. The model takes the news titles corpus as input and utilizes the word-level and title-level, as well as temporal-level, attention mechanism for prediction.
- *LSTM-Tweet*: A similar LSTM model as LSTM-News but replacing the input of news with Trump's tweets.
- *LSTM-Stock*: Taking the stock prices of the 30 comprised blue-chip American corporations as input to evaluate time series prediction, we use an attention-based LSTM model for market trend prediction [31].
- *LSTM-News-Stock*: We concatenate the high-order representations of the LSTM-News and LSTM-Stock as input to fully connected layers to predict the market trend.
- *LSTM-News-Tweet*: We directly concatenate the high-order representation vectors of news and tweets into the BiLSTM network instead of the proposed echoing architecture.
- *NAT Deep Echoing*: To better illustrate the function of the adopted attention mechanism in the word level and title/tweet level, we replace the attention mechanism with an average pooling layer in the news and tweet representation modules and further name it the non-attention (NAT) deep echoing model.
- *NFT Deep Echoing*: We discard the fine-tuning process and directly apply the pre-training model to predict the market trend so that we obtain the non-fine-tuning (NFT) deep echoing model, which is used to demonstrate the effectiveness of the training technique.
- *NTW Deep Echoing*: We remove the temporal weight in the proposed deep echoing modeling, resulting in the non-temporal-weight (NTW) deep echoing model, which can be regarded as an ablation study.
- *Deep Echoing*: The proposed model in this article.

Table 2. Performance of the Proposed Model and Baselines

Methods	Accuracy	MCC
Naive	51.27%	0.025
GBDT	51.66%	0.033
LSTM-Stock	53.02%	0.061
LSTM-News	54.97%	0.1
LSTM-Tweet	54.38%	0.088
LSTM-News-Stock	54.19%	0.084
LSTM-News-Tweet	56.14%	0.123
NAT Deep Echoing	53.99%	0.0825
NFT Deep Echoing	56.53%	0.131
NTW Deep Echoing	58.87%	0.167
Deep Echoing	**60.42%**	**0.206**

6.1.3 Evaluation Metrics. Since the market trend prediction problem is treated as a binary classification problem, we adopt accuracy and the Matthews correlation coefficient (MCC) to measure performance [14]. Furthermore, we also evaluate the profitability of our proposed model through a market trading simulation.

6.2 Market Trend Prediction

6.2.1 Classification Results. To train the proposed model, the length of timesteps used for information summarization is set to be 7, and other parameters including batch size, learning rate, and c are set to be 64, 1e-4, and 0.6, respectively. We compare our approach with the baselines on the binary classification performance. As can be seen from the comparative results shown in Table 2, the last three deep echoing models present the best results compared with other baselines. In particular, the proposed methods that consider both sources of information perform better than other methods with one single data source perspective, which validates our assumption that Trump's tweets together with financial news indeed have stronger power to impact the market in a short time. It is not surprising to see that the traditional prediction methods—Naive Counting and GBDT—almost show no prediction ability due to the highly volatile nature of the stock market. Meanwhile, the attention-based LSTM-Stock model that takes stock prices as input contributes little to market trend prediction, which may be because the market is driven by new information rather than the past and current price. The LSTM-News-Stock model gains an even worse result than the LSTM-News model because of the interference of invalid price information.

In an ablation study, our proposed echoing model also achieves better performance than the LSTM-News-Tweet model, which demonstrates the effectiveness of the proposed echoing module in capturing the interactions between Trump's tweets and relevant news to drive short-term movement. Moreover, the NFT model performs the worst among the three deep echoing models, manifesting the necessity and effectiveness of applying the pre-training and fine-tuning technique to our prediction problem. In addition, by contrasting the two models with and without the temporal weight function, we can find that the full model with temporal weight performs better than the NTW model, which can be partly explained by the existence of the tweet topic drifting phenomenon. As a whole, the contrast results validate the effectiveness of our proposed model.

6.2.2 Market Trading Simulation. To further evaluate the effectiveness of our proposed model intuitively, we conduct a simulated stock trading with a short-term strategy on all 1,183 testing

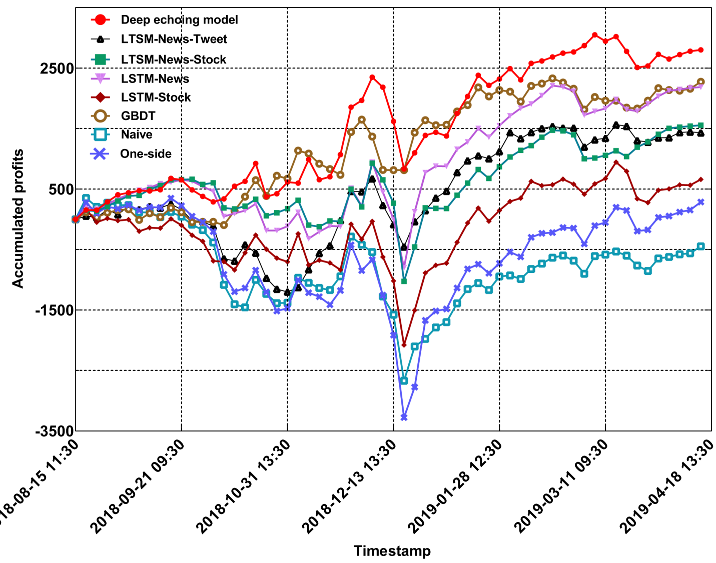

Fig. 4. Accumulated DJIA points curve of different models on the testing dataset.

datasets. We view the closing points of the DJIA as the price of a stock, and our simulation strategy invests the stock based on the predicted trend by the trained model. Specifically, if the future trend is predicted to rise, the simulated trader would buy in and sell it the next day, and thus the gained profit equals the difference of the DJIA index between the two time periods; otherwise, the trader would not take action and the profits stay the same. In this manner, the accumulated profits of the trading simulation can then be calculated based on the prediction results and real-world market movements. For the sake of simplicity, we ignore the transaction cost for each trade. We calculate the simulated accumulated profits gained from all methods with the real DJIA rising from 21,998 points on August 15, 2018, to 26,559 points on April 18, 2019. At the same time, we replace the LSTM-Tweet model with another baseline named *one-side*, which always chooses to buy in considering that a rising trend is observed in the most recent 2 years.

All accumulated profits of various methods are shown in Figure 4. It is clear to see that our proposed deep echoing model outperforms other baselines with a visible gap. It is notable that the accumulated profits with the proposed model achieve 2,800 points in total. Comparatively speaking, the LSTM-News method achieves similar accumulated profits of 2,200 with GBDT but shows a more fluctuating trend. The profit gap between the deep echoing model and LSTM-News illustrates the impact of Trump's tweets and the necessity of taking into consideration the maximum profits. In contrast, the LSTM-Stock method performs much worse than the news-based methods, probably because the EMH of the market and the stock price series offers little information on future trends.

6.3 Parameter Analysis

Considering that the impact of information decreases over time and long-term information may interfere with current information, we also study the influence of the number of timesteps used for information representation on market trend prediction. We vary the length of timesteps T in hours and examine the accuracy of prediction, and the detailed result is presented in Figure 5. We can observe that the model achieves the best performance when timestep $T = 7h$. It is worth noting that six timesteps exist in 1 day (from 9:30 a.m. to 3:30 p.m.) that can be used for prediction, and thus $T = 7$ indicates that the model opportunely utilizes the cross-day data as input for

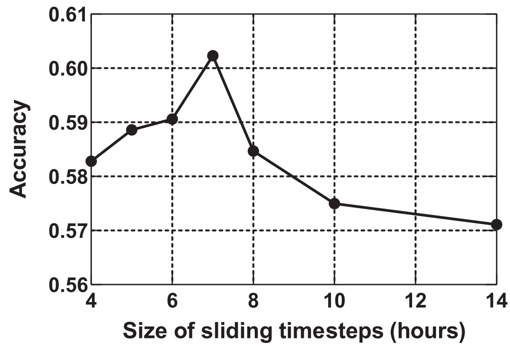

Fig. 5. The influence of timesteps T on accuracy.

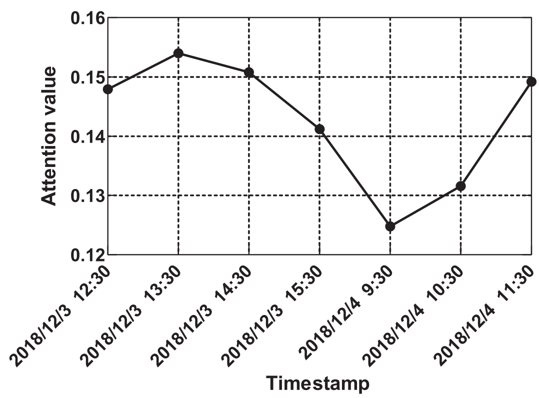

Fig. 6. A case of temporal attention weights for trend prediction at 12:30 on December 4, 2018.

information extraction, whereas when $T \geq 7$, the accuracy decreases as the timesteps become longer. The underlying reason is probably that most relevant news about the market-related tweets is sure to be found in the most recent seven time periods, including the follow-up and previous latest news. Hence, the long timestep cannot provide more information but reduces effective information extraction.

Furthermore, to illustrate the effect of temporal weight attention, we show a case by calculating the attention scores over different timesteps. When we want to predict the future market trend of 12:30 at the current time 11:30 on December 4, 2018, the corresponding temporal weight is calculated and presented in Figure 6. It can be found that the latest information at 11:30 is assigned greater weight than the news reported in several early timesteps. However, some valuable information, such as the reports about U.S.-China friction posted at 13:30 on December 3 still has tremendous impact on the trend prediction with great attention weights, which confirms that the temporal attention mechanism can distinguish market-related information from uninformative information.

6.4 Interpretations from Attention Mechanisms

6.4.1 News Attention. As shown in Table 2, the contrast results of the NAT model and the proposed model demonstrate that the adopted attention mechanism could select market-related keywords to form efficient news representation and further distinguish significant informative news from noisy content. To better illustrate the modeling effect ,we calculate the attention values

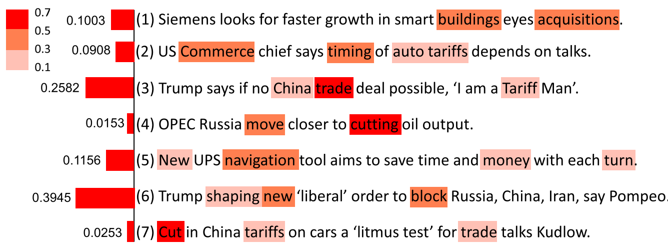

Fig. 7. Marked keywords to represent each news and assigned importance to different news titles.

assigned for each word of the title and every piece of news in each timestep according to the word-level and title-level attention mechanism, respectively. We take the performance of DJIA prediction on December 4, 2018, from 9:30 to 10:30 as an example to show the interpretation ability of our model. It was reported that the DJIA logged its worst day in nearly a month as it declined nearly 600 points, partly due to the trade war.[6] The detailed cases and results of word and title attention are shown in Figure 7.

We mark several words with different colors in seven of the latest news titles according to their assigned weights, where a darker color means larger weights and weights smaller than 0.1 are ignored. As we can see from Figure 7, the word-level attention could highlight informative words from uninformative words. Moreover, the same word in different titles may exert different importance. For instance, when talking about U.S.-China trade friction in the third title, "trade" has the largest weight in the title; however, when the topic shifts toward the tariff cut, "trade" contributes less to the sentence representation in the seventh title.

Furthermore, the histogram in front of each news title indicates its corresponding title-level attention weight. From Figure 7, we can observe that the third and seventh news pieces related to trade friction dominate other information in the current time period. This observation directly affects investor sentiment about the market because Trump's attitude about trade friction expressed in tweets further intensifies the panic of stock market, which also echoes with the stock analysis that the slump is partly attributed to the trade war. Additionally, we also observe that the attention mechanism assigns certain weights to other financial news about Siemens (the first news piece) and UPS (the fifth news piece) to construct a universal market-related information summarization. These cases clearly demonstrate that the news attention mechanism is capable of distinguishing the most impactful news for trend prediction in a chaotic information set.

6.4.2 Tweet Attention. Similarly, we also display the interpretation ability on Trump's tweets. As shown in Figure 8, we take several of Trump's latest tweets posted on December 3, 2018, as an example, in which some words are highlighted with different colors, and a histogram in front of each tweet denotes its weight. According to the results, it is unsurprising that the informative keywords are distinguished among common words, and irrelevant tweets are filtered out with assigning small attention values. For instance, removing tariff is measured as the key point of the second tweet. In contrast, irrelevant tweets including the first and sixth ones contribute less to information summarization, and the second, fourth, and fifth tweets of negotiation with China are assigned the highest importance. This difference indicates that the tweet attention works as a primary filter on all raw data and helps construct a more market-related tweet representation. Obviously, these cases also show the contained rich market-related information in Trump's tweets and explain why the attention mechanism can aid to boost prediction performance.

[6]https://edition.cnn.com/2019/08/14/investing/dow-stock-market-today/index.html.

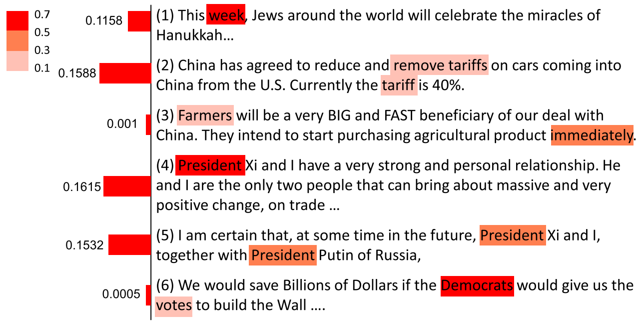

Fig. 8. Marked keywords to represent each tweet and assigned importance to several different tweets.

Table 3. Statistics of Tweets in Relevant and Irrelevant Sets

Category	Topic	Example
Relevant ($p_s \geq 0.7$)	Trade & Economics (55)	Very productive conversations going on with China! We reached a wonderful new Trade Deal with Canada ... A big deal looking good with Mexico!
	Political Events (49)	Saudi Arabia and OPEC will not be cutting oil production We had very substantive negotiations with Kim Jong Un...
	Domestic Events (67)	Americans deserve to know the lowest drug price... I want 5G, and even 6G, technology in the United States...
	Berating & Praising (44)	The New York Times has a new Fake Story that Russians... Our Country is doing so well in so many ways, great jobs...
	Others (20)	Working hard, thank you! VETO!
Irrelevant ($p_s \leq 0.4$)	Trade & Economics (4)	Very good conversation with President Xi Jinping of China ... JOBLESS CLAIMS AT 50 YEAR LOW!
	Political Events (13)	We paid Pakistan Billions of Dollars... Welcome President @AlsisiOfficial of the Arab Republic...
	Domestic Events (14)	President Trump participated in a bilateral meeting... Very proud of the U.S. Senate for voting...
	Berating & Praising (16)	I'm thankful for every day Hillary Clinton is not President! Since my election as President the Dow Jones is up 43%...
	Others (38)	I agree with President Obama 100%! No Collusion - No Obstruction!

Note: All tweets are manually classified into five classes, and the number in the parentheses after each class name denotes the number of tweets belong to this class. We give the detailed definition of the five classes as follows:

- *Trade & Economics:* Includes trade negotiations, trade agreements, and reports about economic conditions, etc.
- *Political Events:* Includes border disputes, meetings with foreign leaders, and international conferences, etc.
- *Domestic Events:* Includes natural disasters, elections, etc.
- *Berating & Praising:* Tweets about berating or praising political leaders or himself.
- *Others:* Irrelevant to the stock market.

6.5 Analysis of Trump's Market-Related Tweets

To investigate the discriminative ability of our model on Trump's tweets and further study whether they always provide market-related information, we explore the correlation of the tweets with the stock market according to Equation (9), which is in the form of a probability p_s. Overall, with respect to the 513 testing samples with Trump's tweets, there are 69.4% of time periods in which the probability exceeds 0.5. The high rate indicates that most of Trump's tweets can be viewed as a steady information source for market prediction.

In addition, we divide the tweets into relevant and irrelevant groups, according to the probability p_s to further analyze which topics of Trump's tweets are most relevant to the market trend. Specifically, the relevant group consists of the tweets with probability higher than 0.7, whereas the irrelevant group is lower than 0.4. For each timestep, we select the tweet with the highest weight assigned by the tweet attention mechanism as the representative tweet of that period. Then we obtain 235 and 85 representative tweets to compose the relevant and irrelevant groups, respectively. The results are presented in Table 3 showing the involved five topics and some corresponding example tweets. Apparently, relevant tweets are about financial and economical events or international politics, especially the trade friction with China, Europe, Mexico, and other market-related information. Comparatively speaking, tweets about entertainment, berating sayings, domestic events, and so forth are considered to have weak correlation with the stock market. To sum up, the results further illustrate the discriminative ability of our model toward Trump's tweets, since it successfully filters out most less informative tweets and contributes to pure market-related information summarization.

7 CONCLUSION

In this article, we study whether and how Trump's tweets impact stock market movement and utilize this new-rising information for more accurate stock prediction. Specifically, we find that the stock market fluctuates heavily when Trump posts tweets, and we further reveal that the echoing reinforced by financial news with impactful Trump tweets sometimes determines the short-term market trend. Then, to model this echoing reinforced impact, we propose a deep echoing model to extract market-related semantic representations from the two data sources, with designed hierarchical attention strategies. Moreover, in the learning phase, we adopt pre-training and fine-tuning techniques and a weighted learning strategy to boost modeling performance. Extensive experiments conducted on the real-world dataset clearly validate our assumption that Trump's tweets can help boost short-term stock market prediction accuracy. The results also confirm the effectiveness of our designed impact process modeling that can result in obvious improvement in terms of the accuracy and produce appreciable profits in a trading simulation compared with several baselines. Furthermore, we find that Trump's tweets about trade and political events are more likely to influence market movements in the short term.

REFERENCES

[1] Jeffrey S. Abarbanell and Brian J. Bushee. 1997. Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research* 35, 1 (1997), 1–24.

[2] Marta Arias, Argimiro Arratia, and Ramon Xuriguera. 2014. Forecasting with Twitter data. *ACM Transactions on Intelligent Systems and Technology* 5, 1 (2014), Article 8, 24 pages.

[3] Wei Bao, Jun Yue, and Yulei Rao. 2017. A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PLOS ONE* 12, 7 (2017), 1–24.

[4] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1 (2011), 1–8.

[5] Francesco Bonchi, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. 2011. Social network analysis and mining for business applications. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011), Article 22, 37 pages.

[6] Jeffery A. Born, David H. Myers, and William J. Clark. 2017. Trump tweets and the efficient market hypothesis. *Algorithmic Finance* 6, 3–4 (2017), 103–109.

[7] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control.* John Wiley & Sons.

[8] Arjun Chatrath, Hong Miao, Sanjay Ramchander, and Sriram Villupuram. 2014. Currency jumps, cojumps and the role of macro news. *Journal of International Money and Finance* 40 (2014), 42–62.
[9] K. Chen, Y. Zhou, and F. Dai. 2015. A LSTM-based method for stock returns prediction: A case study of China stock market. In *Proceedings of the 2015 IEEE International Conference on Big Data (Big Data'15).* 2823–2824.
[10] Yin-Wong Cheung and Lilian K. Ng. 1992. Stock price dynamics and firm size: An empirical investigation. *Journal of Finance* 47, 5 (1992), 1985–1997.
[11] J. Bradford De Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. Noise trader risk in financial markets. *Journal of Political Economy* 98, 4 (1990), 703–738.
[12] Patricia M. Dechow. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18, 1 (1994), 3–42.
[13] Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2014. Using structured events to predict stock price movement: An empirical investigation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP'14).* 1415–1425.
[14] Xiao Ding, Yue Zhang, Ting Liu, and Junwen Duan. 2015. Deep learning for event-driven stock prediction. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'15).* 2327–2333.
[15] Eugene F. Fama. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25, 2 (1970), 383–417.
[16] Qi Ge, Alexander Kurov, and Marketa Halova Wolfe. 2018. Stock market reactions to presidential statements: Evidence from company-specific tweets. *Economics Faculty Scholarship.* Skidmore College.
[17] Rohitha Goonatilake and Susantha Herath. 2007. The volatility of the stock market and news. *International Research Journal of Finance and Economics* 3, 11 (2007), 53–65.
[18] Ziniu Hu, Weiqing Liu, Jiang Bian, Xuanzhe Liu, and Tie-Yan Liu. 2018. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In *Proceedings of the 11th ACM International Conference on Web Search and Data Mining (WSDM'18).* ACM, New York, NY, 261–269.
[19] William Leigh, Russell Purvis, and James M. Ragusa. 2002. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: A case study in romantic decision support. *Decision Support Systems* 32, 4 (2002), 361–377.
[20] Qing Li, Yuanzhu Chen, Li Ling Jiang, Ping Li, and Hsinchun Chen. 2016. A tensor-based information framework for predicting the stock market. *ACM Transactions on Information Systems* 34, 2 (2016), Article 11, 30 pages.
[21] Qing Li, Yan Chen, Jun Wang, Yuanzhu Chen, and Hsinchun Chen. 2017. Web media and stock markets: A survey and future directions from a big data perspective. *IEEE Transactions on Knowledge and Data Engineering* 30, 2 (2017), 381–399.
[22] Qing Li, TieJun Wang, Ping Li, Ling Liu, Qixu Gong, and Yuanzhu Chen. 2014. The effect of news and public mood on stock movements. *Information Sciences* 278 (2014), 826–840.
[23] Manolo Malaver-Vojvodic. 2017. *Measuring the Impact of President Donald Trump's Tweets on the Mexican Peso/U.S. Dollar Exchange Rate.* Master's Thesis. Department of Economics, University of Ottawa.
[24] Hirotaka Mizuno, Michitaka Kosaka, Hiroshi Yajima, and Norihisa Komoda. 1998. Application of neural network to technical analysis of stock market prediction. *Studies in Informatics and Control* 7, 3 (1998), 111–120.
[25] Helen Susannah Moat, Chester Curme, Adam Avakian, Dror Y. Kenett, H. Eugene Stanley, and Tobias Preis. 2013. Quantifying Wikipedia usage patterns before stock market moves. *Social Science Electronic Publishing* 3, 5 (2013), 926–930.
[26] Arman Khadjeh Nassirtoussi, Saeed Aghabozorgi, Teh Ying Wah, and David Chek Ling Ngo. 2014. Text mining for market prediction: A systematic review. *Expert Systems with Applications* 41, 16 (2014), 7653–7670.
[27] John R. Nofsinger. 2005. Social mood and financial economics. *Journal of Behavioral Finance* 6, 3 (2005), 144–160.
[28] Ping-Feng Pai and Chih-Sheng Lin. 2005. A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega* 33, 6 (2005), 497–505.
[29] Tobias Preis, Helen Susannah Moat, and H. Eugene Stanley. 2013. Quantifying trading behavior in financial markets using Google trends. *Scientific Reports* 3 (2013), 1684.
[30] Bhusana Premanode and Chris Toumazou. 2013. Improving prediction of exchange rates using differential EMD. *Expert Systems with Applications* 40, 1 (2013), 377–384.
[31] Yao Qin, Dongjin Song, Haifeng Chen, Wei Cheng, Guofei Jiang, and Garrison Cottrell. 2017. A dual-stage attention-based recurrent neural network for time series prediction. arxiv:cs.LG/1704.02971
[32] Radim Řehůřek and Petr Sojka. 2010. Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks.* 45–50.
[33] Eduardo J. Ruiz, Vagelis Hristidis, Carlos Castillo, Aristides Gionis, and Alejandro Jaimes. 2012. Correlating financial time series with micro-blogging activity. In *Proceedings of the 5th ACM International Conference on Web Search and Data Mining (WSDM'12).* ACM, New York, NY, 513–522.

[34] Robert P. Schumaker and Hsinchun Chen. 2009. Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Transactions on Information Systems* 27, 2 (2009), Article 12, 19 pages.
[35] Lei Shi, Zhiyang Teng, Le Wang, Yue Zhang, and Alexander Binder. 2019. DeepClue: Visual interpretation of text-based deep stock prediction. *IEEE Transactions on Knowledge and Data Engineering* 31, 6 (2019), 1094–1108.
[36] Andrei Shleifer and Robert W. Vishny. 1997. The limits of arbitrage. *Journal of Finance* 52, 1 (1997), 35–55.
[37] Jianfeng Si, Arjun Mukherjee, Bing Liu, Sinno Jialin Pan, Qing Li, and Huayi Li. 2014. Exploiting social relations and sentiment for stock prediction. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP'14)*. 1139–1145.
[38] Paul C. Tetlock. 2007. Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance* 62, 3 (2007), 1139–1168.
[39] Dat Thanh Tran, Alexandros Iosifidis, Juho Kanniainen, and Moncef Gabbouj. 2018. Temporal attention-augmented bilinear network for financial time-series data analysis. *IEEE Transactions on Neural Networks and Learning Systems* 30, 5 (2018), 1407–1418.
[40] Yang Yu, Wenjing Duan, and Qing Cao. 2013. The impact of social and conventional media on firm equity value: A sentiment analysis approach. *Decision Support Systems* 55, 4 (2013), 919–926.

Received December 2019; revised April 2020; accepted May 2020