Research Article

Anticipating Stock Market of the Renowned Companies: A Knowledge Graph Approach

Yang Liu ⓘ,[1] Qingguo Zeng ⓘ,[2] Joaquín Ordieres Meré ⓘ,[1] and Huanrui Yang[3]

[1]*Department of Industrial Engineering, Business Administration and Statistics, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, Madrid 28006, Spain*
[2]*School of Mathematical Sciences, South China Normal University, Guangzhou 510631, China*
[3]*Electrical and Computer Engineering Department, Duke University, Durham 27708, USA*

Correspondence should be addressed to Yang Liu; yang.liu00@alumnos.upm.es

An increasing number of the renowned company's investors are turning attention to stock prediction in the search for new efficient ways of hypothesizing about markets through the application of behavioral finance. Accordingly, research on stock prediction is becoming a popular direction in academia and industry. In this study, the goal is to establish a model for predicting stock price movement through knowledge graph from the financial news of the renowned companies. In contrast to traditional methods of stock prediction, our approach considers the effects of event tuple characteristics on stocks on the basis of knowledge graph and deep learning. The proposed model and other feature selection models were used to perform feature extraction on the websites of Thomson Reuters and Cable News Network. Numerous experiments were conducted to derive evidence of the effectiveness of knowledge graph embedding for classification tasks in stock prediction. A comparison of the average accuracy with which the same feature combinations were extracted over six stocks indicated that the proposed method achieves better performance than that exhibited by an approach that uses only stock data, a bag-of-words method, and convolutional neural network. Our work highlights the usefulness of knowledge graph in implementing business activities and helping practitioners and managers make business decisions.

1. Introduction

Research on stock market of company prediction has become popular with the introduction of Fama's efficient market hypothesis (EMH) [1], but in recent years, an increasing number of people have found that stock market of the renowned company changes is random, complex, and unstable given that they are affected by many factors. Stock market of company changes is affected by cultural aspects that fundamentally influence investor sentiment; traditional culture often determines people's investment psychology, which in turn, affects the inflow and outflow of stock market capital. Company-related issues, such as the effects of company operations on stock prices and market factors [2]. Using these factors as bases for accurately predicting stock price movement gives investors hope that the maximum profits will be achieved with the least stock investment. Correspondingly,

analyzing stock market movements is both challenging and attractive to researchers and investors. Moreover, as the development of knowledge graph in natural language processing (NLP), researchers in the financial field began paying attention to text mining in financial news. As maintained by EMH and artificial intelligence technique, the best results are achieved when news information is used to forecast stock market movement, which can also control financial risk in business activities.

In May 2012, Google formally announced its knowledge graph project, which is aimed at improving the search engines effectiveness and search quality as well as user experience with the engine [3]. The development of knowledge graph promoted the extensive use of artificial intelligence technology in smart search, smart question-and-answer tasks, and intelligent finance. In finance, knowledge graph is designed to find relationships amongst entities, such as the management

of companies, news events, and user preferences. These entities can be used by investors to achieve efficient financial data-based decision-making and obtain business insights into predicting stock market [4]. For these reasons, the current research focused on how to use knowledge graph to improve the accuracy of stock price forecasts. Knowledge graph is databases that implement semantic searches by preserving relationships amongst multiple entities. Based on event tuples in knowledge graph [5], we can infer the definition of an event tuple as *(A, P, O)*, where *A* represents an agent, *P* denotes a predicate, and *O* represents an object [6]. In view of the objects converted by each event tuple are known, when they are converted to vectors, which loses more semantic information. Event tuples link relevant elements together and can be used as an efficient method of improving predictive accuracy. As a result, we propose a learning model of event tuples, which retains their semantic features at the maximum.

Knowledge graph embedding is a kind of representation learning in knowledge graph. Currently, several algorithms [3] explore the mapping relationship between entities and relationships in the translation distance model. The TransE model [7] is a computationally efficient predictive model that satisfactorily represents a one-to-one type of relationship. It has been used as a basis in the development of different translation distance models. The TransH model [8] maps the head and tail of a vector onto a hyperplane, after which these two components complete the translation process on the hyperplane. The TransR model [9], which is based on the TransE model, consists of entities and relationships that are located in different dimensions of space. For each relationship, the model defines a matrix that is used to convert entity vectors into space. The model is also dynamically determined by an entity relationship. The TransD model [10] is an improvement of the TransR model, which the former uses as a basis in considering a transformation matrix. In comparing these models, solving the head and tail of a news tuple for mapping in a single space necessitates the simple mapping of one-to-one entity extraction features. This functionality is found in the TransE model; this study is used to establish a feature combination model for news event metagroups.

This work selects six globally renowned companies to forecast their stock price movement, namely, Apple, Microsoft, Samsung, Boeing, Google, and Walmart as application scenarios. According to different data sources between Thomson Reuters and Cable News Network, we reveal the result of the case studies in some algorithms, which illustrates that the performance of combined feature outperforms that using only stock data, using a bag-of-words algorithm, and using convolutional neural network. Although there are already some examples of the powerful application of deep learning in NLP [11], such as speech recognition, text classification, and machine translation, most of the previous research in predicting stock price movement [12] is based on the semantic information news, which ignores the semantic features of structured events. The application of deep learning and knowledge graph on the renowned companies' stocks is rarely available. Thus, our work provides a viable application framework for financial markets, which can also be extended to other aspects of the financial filed.

For stock market prediction, we formulated a knowledge graph embedding-driven approach that involves four main phases (Figure 1). The first is data retrieval, which begins with searching keywords such as "Apple" or "Google" over the Thomson Reuters or Cable News Network (CNN), as shown in Figure 2. The yellow part is the headline of financial news and the orange part is the release time of the financial news. A web crawler is then used to obtain financial news from the sites and match them with corresponding stock data, combining all this information into one corpus. The second phase involves preprocessing, which encompasses corpus analysis, text normalization, word tokenization, label tagging, and word-to-vector transformation. The features derived from word embeddings and stock data layers are then selected for the computation of eigenvalues, after which a feature vector is generated using machine learning and knowledge graph embedding. The third phase is model creation, in which stock market prediction labels (increase or decrease) are assigned to financial news to training a classification model. Finally, the finance decision relies on the predictive performance of this framework. Specifically, a small improvement directly affects decision-making, which increases the renowned company's profits. The fourth phase involves model evaluation, wherein the results and the extraction of conclusions of each machine learning model are analyzed.

The remaining part of the paper is organized as follows. Section 2 is a review of existing studies on the use of machine learning in stock market prediction. Section 3 introduces the methodology adopted in this study, including the data description and feature selection approach. Section 4 presents the classification results, and Section 5 discusses the results and implications. Section 6 concludes with the summary.

2. Literature Review

In Table 1, the application of machine learning techniques to stock market prediction is an emerging research field. The advantage of machine learning models is their ability to facilitate the processing of large amounts of data. A procedure common to these models is the concatenation of features from different sources into a feature vector. For most machine learning models used, research is focused on predicting stock trends (increase or decrease).

Various feature extractions have already been proposed to predict the stock price of the renowned companies. In the early years, most of the company's stock prediction depends on the empirical analysis of the econometric model, namely, the feature extracted is the raw data of the company's stock. However, the previous work ignored the potential impact of unstructured data on company stocks. Since the bag-of-words is widely utilized in the task of document classification, the frequency of word appearance can be used as a feature of the classification. Hence, the bags of words model only calculate the frequency of words and do not consider the word order or word sparsity in the context, which directly impacts the prediction result. Furthermore, EMH had found that the emotional impulses of the renowned company investors often observe abnormal fluctuations in company stocks.
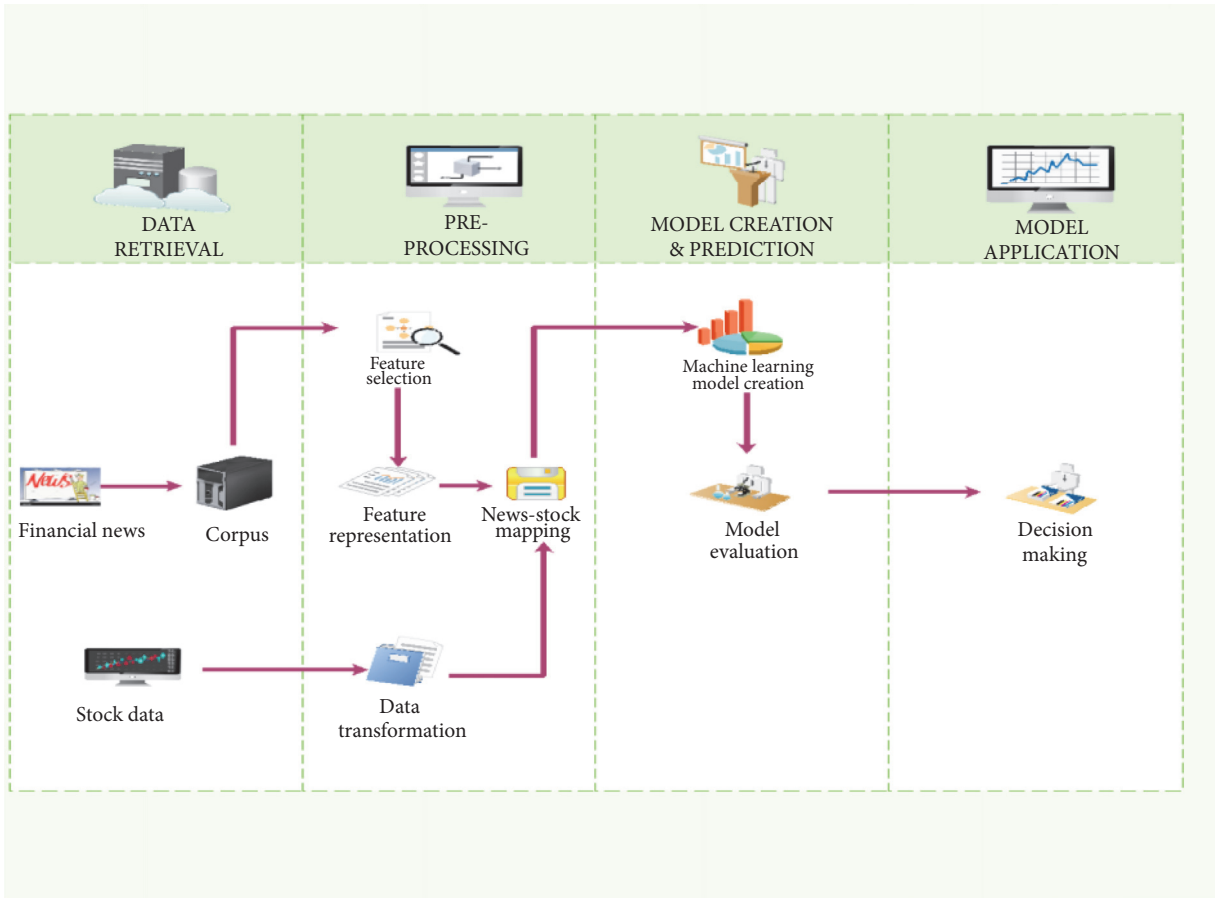
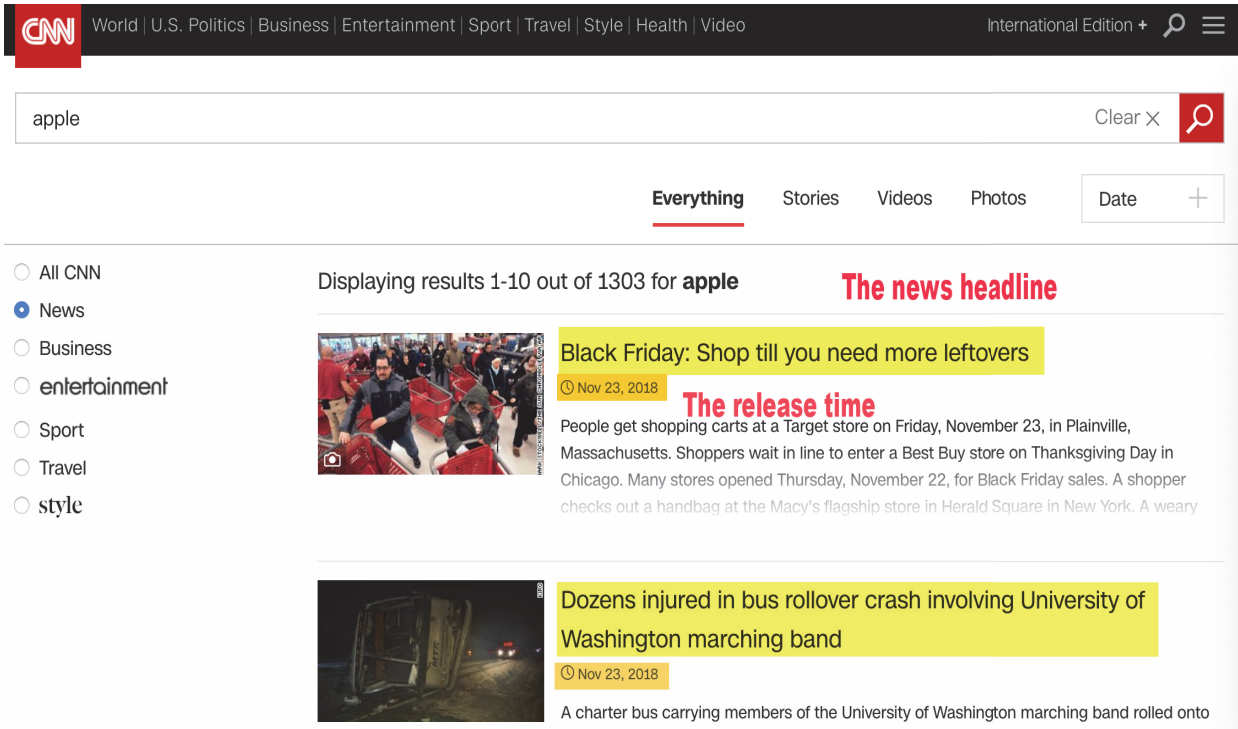FIGURE 1: Research workflow for machine learning and knowledge graph.

FIGURE 2: Example of collected financial news in CNN.

Table 1: A review of financial expert system. ANN, NB, KNN, SVM, DT and SVR stand for artificial neural network, naive Bayes, k nearest neighbors, support vector machine, decision tree, and support vector regression, respectively. Measure metrics: mean squared error (MSE); root mean square error (RMSE); area under the curve (AUC).

Paper	Text type	Feature extraction	Model type	Performance measure
[13]	Professional newspaper	Stock data	NB	Accuracy
[14]	Professional newspaper	Stock data	SVM	Accuracy
[15]	Message board	Stock data	Regression	Correlation coefficient
[16]	PRNewswire	Stock data	KNN, SVM	Accuracy, F1-score
[17]	Aggregate news	Stock data, bag-of-words	SVR	MSE, accuracy
[18]	Corporate disclosures	Stock data, bag-of-words	Multiple	Accuracy, recall, precision, F1-score
[19]	Financial news	Stock data, sentiment	DT	Accuracy, recall, precision, F1-score
[20]	Web media	Stock data, bag-of-words.	SVR	Accuracy, recall, precision, F1-score
[21]	Mixed type	Stock data, bag-of-words sentiment,	SVR	RMSE, accuracy, precision
[22]	Message board	Stock data, sentiment	SVM	Accuracy
[23]	Mixed type	Stock data, technical indicators, sentiment	ANN, SVM, DT	Accuracy, AUC, F1-score, precision
[24]	None	Stock data	Regression	Correlation coefficient
[25]	Media news	Stock data	Regression	Correlation coefficient
Our financial expert system	Financial news	Stock data, technical indicators, bag-of-words, syntax	Multiple	Accuracy, F1-score

Tetlock [26] adopted the popular news from the Wall Street Journal to find that news sentiment has predictive power on company stocks. Chen et al. [27] found information comentions have a significant influence on the stock return through sentiment analysis. Furthermore, the investors' sentiment trends after the positive news, which results in a buying trend, and higher stock market prices while after negative news stocks are sold result in a decrease of the price. However, sentiment analysis can only be employed in specific texts. If sentiment is implicit rather than direct emotional words, then the ability of using sentiment analysis for predicting the stock price of the renowned company is relatively limited. Mittermayer and Knolmayer [16] illustrated that news-CATS achieves a performance that is superior to that of other ATC prototypes used to forecast stock price trends. Li et al. [21] proposed a media-aware quantitative trading strategy by using sentiment information on web media. The simulation trading return was up to 166.11%. Nguyen et al. [22] proposed a feature topic sentiment to improve the performance of stock market prediction. Their method achieved an accuracy that is 9.83% better than that of the historical price method and 3.03% better than that of the human sentiment method.

We utilize the characteristics of syntax analysis as being proposed in [6, 12]; namely, a structured tuple is extracted from an unstructured text based on the semantic structure of each piece of news. Knowledge graph can enrich the structured representation of the news event and effectively retain feature vectors for the news event. The main feature extraction in the previous studies [28, 29] is sentiment analysis, which neglected the event characteristics in the text. Furthermore, the existing literature [23, 29] had proved the positive effect of technical indicators on stock market prediction. In summary, our research highlights syntax analysis in financial news, which also incorporates with other features extraction (stock data, technical indicators, and bag-of-words). Because of the variety of considered features, this research will deliver an improved prediction of the stock market value for renowned companies in at least 3.6%.

Previous research [30] was applied with the traditional machine learning algorithms, especially, since the powerful application ability of deep learning based on machine learning. Deep learning is utilized in several studies for predicting stock price movement. Kraus and Feuerriegel [31] forecasted the stock returns based on financial disclosures,

Table 2: Quotes and renowned company names.

Financial news dataset	Stocks	Company names	Time interval	Number of news headlines
Thomson Reuters	AAAL	Apple Inc.	3.10.2011- - -31.7.2017	6423
	MSFT	Microsoft Corporation	30.9.2011- - -29.11.2017	6623
	SSNLF	Samsung Electronics Co., Ltd.	30.9.2011- - -29.11.2017	6900
CNN	BA	The Boeing Company	8.9.1998- - -23.2.2017	2999
	GOOG	Google Inc.	25.5.2006- - -24.2.2017	5278
	WMT	Walmart Inc.	7.5.1997- - -24.2.2017	4913

Table 3: Samples extracted from the dataset.

Date	News headline	News headline event tuple
3/18/2012	Chinese writers' group sues Apple: state media	Chinese writers' group, sues, Apple
4/13/2012	German court upholds ban on Apple push email	German court, upholds ban on, Apple push email
4/4/2012	Apple, two publishers hold out against settlement-report	Apple- two publishers, hold out against, settlement-report
3/26/2012	Apple CEO Cook meets Chinese officials	Apple CEO Cook, meets, Chinese officials
4/25/2012	US STOCKS-Apple results put Nasdaq on track for best gain of year	Apple, results, Nasdaq on track for best gain

and their result demonstrated that a higher directional of deep learning surpasses traditional machine learning. Ding et al. [5] illustrated that deep learning also can forecast stock market of the renowned company. Sim et al. [32] proposed that technical indicators transform into images of the time series graph, which examines the applicability of deep learning in the stock market. Overall, we adopt the multiple models for prediction stock price of the renowned company, which proves the reliability of our proposed model as comparing different algorithms. This work applies deep learning incorporated with knowledge graph embedding for feature extraction, which examines the applicability of combined features methods in the renowned company stock price movement.

3. Materials and Methods

We developed a knowledge graph-based approach that consists of three steps, namely, data description, data preprocessing, and feature selection.

3.1. Dataset Description. Table 2 shows the custom financial news corpus built with headlines from two datasets. The first dataset contains news articles published by Thomson Reuters, including those regarding Apple Inc. (AAPL), Microsoft Corporation (MSFT), and Samsung Electronics Co., Ltd. (SSNLF). The second dataset comprises news reports published in CNN, including reports on the Boeing Company (BA), Google Inc. (GOOG), and Walmart Inc. (WMT). It also consists of financial news headlines published at specific time intervals, with each news report accompanied by a title and a release date. Titles are used for event embedding and feature extraction, and release dates are used as a reference in ensuring alignment between corresponding financial news

and trading data from a time series. As shown in previous work [5, 31], using a headline to build corpora can help reduce noise in text mining as headlines concisely represent the content of a text. We used only the news headlines from Reuters and CNN for the prediction of stock price movement.

Daily stock data from index report in each company are collected by Yahoo Finance in the same period during stock data and financial news headlines. Daily trading data, which are common predictors of stock price [23, 33], and technical indicator features were used in our model. There are opening price, closing price, high price, low price, and volume and three technical indicators.

Table 3 shows some examples of filtered financial news headlines. To illustrate, 6423 headlines regarding Apple Inc. were extracted and then reduced to 2799 headlines after filtering via Reverb [34]. Let us take the article "*What is next for Apple's board*?", published on 6 October 2011, as a specific example. The title of the article cannot be transformed into an event tuple using Reverb. Because the sentence is in the interrogative form, there is no event tuple that constitutes a rule. After the matching of time-wise data and stock data, 941 headlines were left. Daily news and stock data were aligned to create input-output pairs, except the days when no news was released. In 9 January 2012, for instance, three news articles were reported, but we chose only one headline for alignment with stock market data. News events possibly happen several times within one day, but they do not happen every day, unlike stock trading, which happens daily, except on nontrading days falling on weekends or holidays.

Table 4 shows that the matches found are pairs between event tuples and stock data. From this complete dataset, we used 80% of the samples in training data and the remaining 20% for testing data. This selection method is the same as the previous literature [5, 24].

TABLE 4: Training and testing datasets.

Data source	Company	Training	Testing	Total
Reuters	Apple	780	161	941
	Microsoft	832	208	1040
	Samsung	855	214	1069
CNN	Boeing	528	132	660
	Google	1010	252	1262
	Walmart	798	200	998

TABLE 5: Number of event characteristic labels for each company.

	Companies					
Label	Apple	Microsoft	Samsung	Boeing	Google	Walmart
0	502	967	1433	133	661	206
1	537	432	379	200	245	287
2	661	557	378	152	389	233
3	692	707	724	233	683	438
4	407	228	193	147	196	180
Total	2799	2891	3107	865	2174	1344

TABLE 6: Features of the prediction model.

	Feature	Feature expression
1	Stock price yesterday + yesterday's trend (in Section 3.3.1)	$price_{t-1}$, $t\_indicator_{t-1}$
2	Same + bag-of-words (in Section 3.3.2)	$price_{t-1}$, $t\_indicator_{t-1}$, tf_idf_{t-1}
3	Same + some features inferred through a convolutional neural network (in Section 3.3.3)	$price_{t-1}$, $t\_indicator_{t-1}$, cnn_{t-1}
4	Same + some features inferred through feature combination (in Section 3.3.4)	$price_{t-1}$, $t\_indicator_{t-1}$, $f\_combine_{t-1}$

3.2. Data Preprocessing. The following three steps are for preprocessing, which prepares for feature extraction and model creation.

(1) Tagging the label for each news headline. Five possible labels are characterized by categorical values as follows: 0 for an extremely negative label, 1 for a negative label, 2 to signify a neutral label, 3 for a positive label, and 4 to denote an extremely positive label. According to the time characteristics of each news headline, an event is manually tagged with the label for each news headline. Table 5 shows the labeling applied for each company: Label 0 means that a company's competitor has happened in this event; label 1 means that the company lost something; label 2 means that it did not cause any impact on the company; label 3 means that this event enabled the company to obtain something; and label 4 means that the company increased its profits or created more value.

(2) Word vector transformation. We used the word2vec [35] algorithm to train word embedding and set the number of dimensions to 300. The word embedding was also trained using the Google News dataset, which contains 100 billion words that are characterized by a continuous bag-of-words structure.

(3) Technical indicators calculation. Three additional technical indicators that are calculated on the basis of daily trading data were used as follows:

(a) Stochastic oscillator (%K). This indicator is a momentum analysis method created by George C. Lane. When the price trend rises, the closing price tends to approach the highest price of the day. When the price trend declines, the closing price tends to approach the lowest price of the day [36].

(b) Larry Williamss %R indicator. This indicator is an oscillation indicator that measures the ratio of the highest price to the daily closing price. It indicates the proportion of stock price fluctuations in a certain period, thereby providing a signal out of the reversal of a stock market trend [37].

(c) Relative strength index (RSI). Buying and selling intentions in the market are analyzed by comparing the closing prices in a given period. Stocks that have had more or stronger positive changes have a higher RSI than do those that have had more or stronger negative changes. Strength indicators fall between 0 and 100; investors sell if this value is ≥ 80 and buy if it is ≤ 20 [23, 36].

3.3. Variable/ Feature Selection. To assess the effectiveness of applying the prediction model on the basis of financial news, we designed four sets of features for predicting stock price movement (Table 6). Features 3 and 4 are used for event characteristics. Each feature is explained in the succeeding subsections. The target output consists of a binary variable, for which a value of 1 indicates that the closing price at day $t + 1$ will be higher than that at day t, and a value of 0 indicates that the closing price at day $t + 1$ will be lower than that at day t.
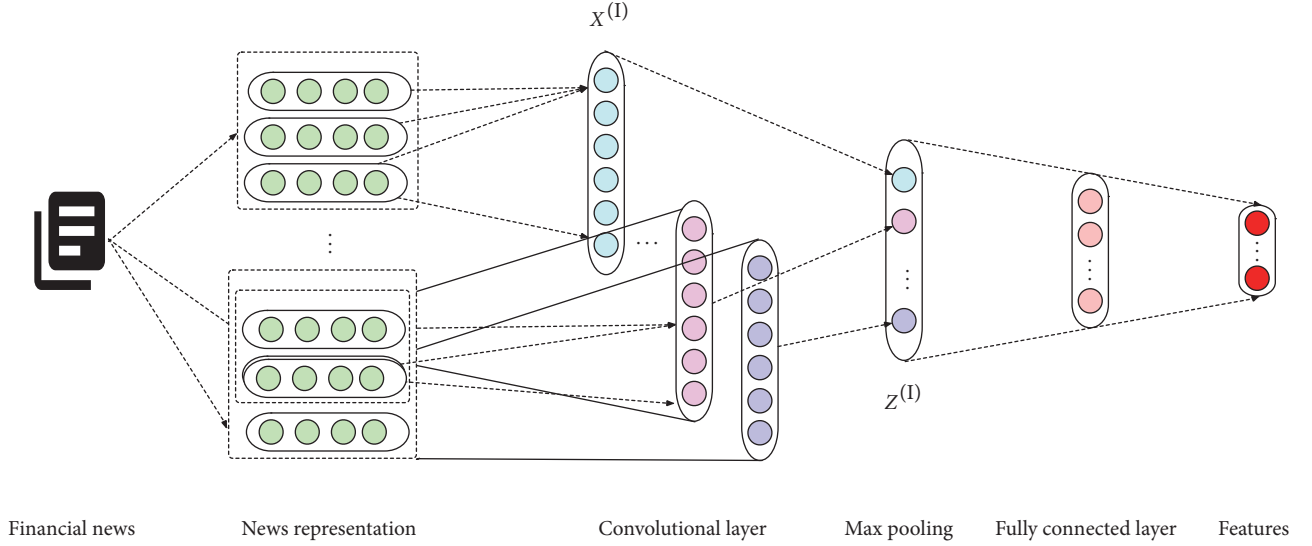
FIGURE 3: Convolutional neural network architecture proposed for feature selection.

3.3.1. Stock Data Only. We considered historical price as the input for predicting stock price movement and used it as a baseline for comparison with other sets of features. The features used to train the machine learning model that uses only stock data are $price_{t-1}$ and $t\_indicator_{t-1}$. The output is the indicator of price movement (increase or decrease) at each transaction date examined.

3.3.2. Bag of Words. The input feature set is based on a bag of words for the news and price trend for stock. Previous studies [18, 38, 39] have widely used and confirmed the feasibility of the bag-of-words algorithm, but this method disregards elements such as grammar and word order in the text. In the present study, we first prepared each headline for data preprocessing and then transformed the preprocessed headline into a feature vector using the term frequency-inverse document frequency (TF-IDF) algorithm [40], which assigns high weight to eigenvectors. The studies [41, 42] have strongly proven the effectiveness of the TF-IDF algorithm in feature extraction from news headlines. It estimates the frequency of a term in one document over the maximum in a collection of documents and assesses the importance of a word in one set of documents. Such importance increases proportionally with the number of word appearances in a document. The features used to train the bag-of-words machine learning model are $price_{t-1}$, $t\_indicator_{t-1}$, and tf_idf_{t-1}, which are the price movements (increase or decrease) at each transaction date examined.

3.3.3. Convolutional Neural Network. Given the sequence of words in a headline, the word2vec model [35] can be used to embed each of these words in a real valued-vector x. In this work, we set the dimension n of each word vector at 30 (i.e., $x \in \mathbb{R}^{30}$). We concatenated the word vectors of all the words in a headline sequentially to form a matrix $X = [x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(4)}]$ [35] as the input to a convolutional neural network model. For a headline with N words, the resultant input matrix has dimensions of $30 \times N$, and the dimension of news representation also is 30.

In Figure 3, this convolutional neural network model is made up of four consecutive layers: the first layer is the input, the second layer is convolutional, the third layer is a max-pooling, and the last layer is a fully connected layer. The convolutional and max-pooling layers were designed using the text-attentional convolutional neural network [43], which effectively carries out sentiment classification. In the convolutional layer, input matrix X convolves with a kernel $w \in \mathbb{R}^{n \times k}$, where n is the size of a word vector (30 in our work), and its dimension is 50. k denotes the size of a sliding window ($k=3$ in this study). The computation can be formulated as follows:

$$Z_i = \sigma\left(w X_{i:i+k-1} + b\right), \tag{1}$$

where $X_{i:i+k-1} = [x_i, x_{i+1}, \ldots, x_{i+k-1}]$ is the portion of input X falling within the sliding window, b denotes the optional offset, and $\sigma(\cdot)$ is the sigmoid function.

In the next step, we used the pooling layer to reduce the convolutional neural networks parameter space and consequently minimize information loss in the pooling processing and capture the most important features. The eigenvector of the filter in the pooling layer is cut into 3 chunks and a maximum value is taken in each chunk; we obtained 3 eigenvalues. Convolutional output vector Z was split into p windows, and only the maximum feature as 3 in each window was kept for passage onto the final fully connected layer. The fully connected layer is linear regression, and the output layer is the feature classification between 0 and 1.

Our convolutional neural network model is intended to extract a feature vector from the fully connected layer to describe the emotional characteristics of input headlines. Following the intuition that a good feature vector should lead to the accurate classification of the event characteristics of headlines, we attached a softmax layer after the fully
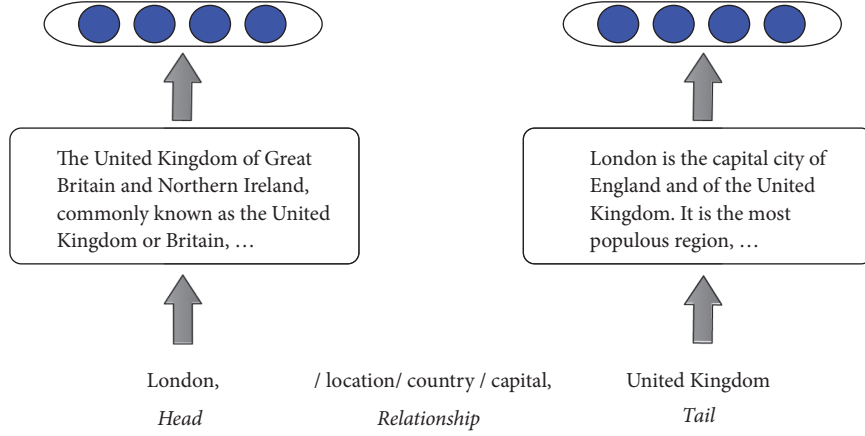
FIGURE 4: Text description of the entity.

connected layer when the convolutional neural network model was trained. The entire model was trained to classify the five emotional labels that describe event characteristics, and the resultant model will be able to provide a meaningful emotional feature vector for each inputted headline. Meanwhile, we define the loss function of the convolutional neural network model as E_d. The features used to train this machine learning model are $price_{t-1}$, $t\_indicator_{t-1}$, and cnn_{t-1}, which are the price movements (increase or decrease) at corresponding transaction date.

3.3.4. Combined Features. In the introduction section of this paper, we described various translation models. Because the TransE model represents a one-to-one relationship between two entities [7], the relationship amongst many entities must be continuously incorporated into texts and knowledge graph [44]. This model combines a convolutional neural network with textual information extraction, which fully exploits the semantic information in a knowledge graph and text [45, 46]. In Figure 4, knowledge graph contains rich semantics in entity description texts, but it is not fully utilized in feature extraction. Most of the existing text representation models simply train the text into word vectors through word2vec, which obtains the text representation by means of averaging, etc. Hence, these ways often lose more semantic information. We therefore proposed to extract feature vectors from news texts using a convolutional neural network model combined with the TransE model, which also fully integrates the two parts of the feature information.

In Figure 5, this architecture used in the feature combination model encompasses two parts, namely, a convolutional neural network (Section 3.3.3) and the TransE model. The feature combination sets are the average of each word vector in an entity and it was obtained using the word2vec model. The two entity vectors are mapped into the same relational space, and these entities using a trained low-rank weight matrix [47].

As assumed in the TransE model, relationship vector R should satisfy "$E + R \approx T$". The model can be represented as a parameter set $\emptyset = (L_r, R_r, \theta)$, where X, E, and R

represent a word, an entity, and a relation, respectively. And r are the mapping matrices for entities in the structural model and represents the weights of the convolutional neural network. For instance, the sentence is "*Samsung sues Apple for infringement*", so "*Samsung + Apple ≈ sues*". The loss function E_s of this structural model is defined as follows:

$$E_s = \|h + r - t\|_2^2, \tag{2}$$

where h, r, and t represent head entity, relationship, and tail entity in event tuple, respectively [47].

And text representation E_d is consistent with Section 3.3.3, which denotes:

$$E_d = f_r(h_s, t_d) + f_r(h_d, t_s) + f_r(h_d, t_d), \tag{3}$$

where $f_r(h_d, t_d) = \|h_d + r - t_d\|_2^2$ represents the head entity and the tail entity is in text representation. $f_r(h_d, t_s) = \|h_d + r - t_s\|_2^2$ and $f_r(h_s, t_d) = \|h_s + r - t_d\|_2^2$ represent one of the head entity h and the tail entity t in text representation, another one is in structure representation.

Moreover, we combined two types of representation learning (convolutional neural network and feature combination) to map news titles; these vectors will be into feature vectors. The relationship vector R identical to the result of the feature extraction layer of the convolutional neural network in the structural model. Besides, we combined this loss with the classification loss of the convolutional neural network using *L2* regularization, which obtains the overall loss function E for feature selection, that is,

$$E = E_s + \alpha E_d + \beta \|\emptyset\|^2, \tag{4}$$

where $\|\emptyset\|^2$ is a regular item. α and β are hyperparameters, which measure the loss of text information and the weight of the regular items, respectively. W denotes the convolutional kernels in the layer and T' is a negative sample set of T [7].

$$T' = \left\{ \left(h', r, t\right) \mid h' \in E \right\} \cup \left\{ \left(h, r, t'\right) \mid t' \in E \right\} \\ \cup \left\{ \left(h, r', t\right) \mid r' \in R \right\}, \tag{5}$$
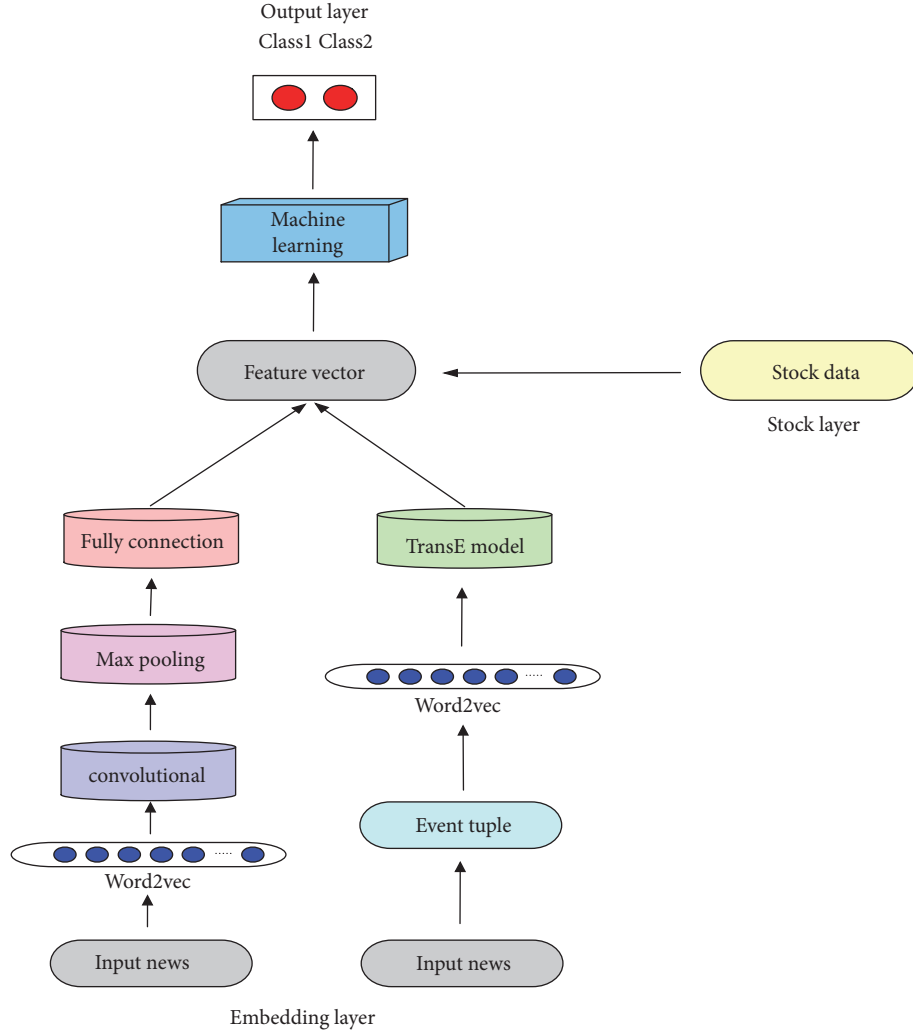
FIGURE 5: Proposed architecture for the feature combination model.

where the head entity and tail entity are randomly replaced by an entity or relationship in another tuple. In particular, if the replaced tuple is already in T, it would not be added to the negative sample. Since both h and t represent two types of entities, there are structure-based representations and text-based representations in interval-based loss functions. The stochastic gradient descent (SGD) is employed to minimize the above loss function.

In this structure, we choose the optimal parameters as followe: the learning rate of SGD $\lambda = 0.001$, the dimension of the representation vector of entities and relationships $k = 100$, the word vector dimension of the entity description text $n = 100$, and batch size during training is 1440. Convolution layer window size is $W \in \{1, 2, 3\}$. This experiment is performed in 1000 iterations of training and the optimal parameters are based on the testing set.

Hence, the features used for training the machine learning model are $price_{t-1}$, $t\_indicator_{t-1}$, and $f\_combine_{t-1}$, which are the price movement (increase or decrease) at corresponding transaction date.

4. Experiments and Results

4.1. Experimental Setting. This section compares the performance of different models in feature extraction from financial news expended with full sets of features. The evaluation was intended to obtain evidence that the proposed feature combination model is superior to other feature selection models in predicting the stock price movement. As shown in Table 7, we chose linear models, for instance, logistical regression and naive Bayes. Others are nonlinear models, for instance, the ensemble learning (random forest, adaboost, and gradient boosting) for the comparison. We constantly adjusted the parameters in the grid search and selected the optimal parameter value; their parameter value has proven to work well on machine learning method [29, 48].

We used a computer consisting of an Intel Core i5 processor with four cores running at 2.9 GHz and 8 GB RAM under the MacOS platform. We used the Scikit-learn library in Python in the experiments involving traditional machine learning algorithms and TensorFlow 1.4 in the

TABLE 7: List of machine learning methods.

Method	Classification	Parameter and their value
Traditional machine learning	Decision tree (DT)	Criterion= gini, splitter= best, max depth=None, min samples split=2, min samples leaf =1, min weight fraction leaf=0.0,
	Logistical regression (LR)	Random state=1
	Naive Bayes (NB)	No kernel estimator
	Stochastic gradient descent (SGD)	Loss=hinge, penalty=l2
	Support vector machine (SVM)	Polynomial kernel function with exponent = {1, 2}, RBF kernel function with gamma = 0.01
Ensemble learning	Random forest (RF)	Randomly sampled as candidates at each split = log2 + 1, max-depth=2, random state=0
	AdaBoost (AB)	Polynomial kernel function with exponent = {1, 2}, RBF kernel function with gamma = 0.01, n- estimators=100
	Gradient boosting (GB)	N-estimators =100, learning rate =1.0, max-depth=1, random state=0

TABLE 8: Abbreviations of the results for each algorithm considered.

Algorithms	Stock data (SD)	Bag of words (BOW)	Convolutional neural network	Feature combination model (FC)
DT	DT_1	DT_2	DT_3	DT_4
LR	LR_1	LR_2	LR_3	LR_4
NB	NB_1	NB_2	NB_3	NB_4
SGD	SGD_1	SGD_2	SGD_3	SGD_4
SVM	SVM_1	SVM_2	SVM_3	SVM_4
RF	RF_1	RF_2	RF_3	RF_4
AB	AB_1	AB_2	AB_3	AB_4
GB	GB_1	GB_2	GB_3	GB_4

experiments involving deep learning and the TransE model. The abbreviations used for comparing classification performance are presented in Table 8. During testing, 2-fold cross-validation is applied to evaluate the stability of the models. And we compared performance in predicting stock price movement for the next day with a test dataset and evaluated the performance of the models in terms of accuracy and F1-score [49].

4.2. Results

4.2.1. Thomson Reuters Case Study. Thomson Reuters is a multinational mass media and information firm. Three renowned companies on the website were selected for the analysis. The first is Apple Inc., which is an American multinational technology company that designs, develops, and sells consumer electronics, computer software, and online services. The second company selected is Microsoft, which is an American multinational technology company that develops, manufactures, supports, and sells computer software, personal computers, and other services. The last company chosen is Samsung, which is a South Korean multinational electronics company. We considered data from these three typical technology companies for our problem.

Measurements of average accuracy are shown in Table 9, and the results for each renowned company with respect to prediction based on different features are illustrated in Figure 6. We calculated the average of the four feature construction approaches for each model for comparison amongst the three companies. The proposed feature combination model generated the best results, achieving average accuracy levels of 61.63%, 59.18%, and 58.48% for Apple, Microsoft, and Samsung, respectively. These figures are consistent with the rates reported in previous research. However, many studies analyzed only one company or used only one algorithm. The current research probed into three companies on the basis of information from a common data source to build a prediction model with different feature selection functionalities and different algorithms. We used stock data, a bag-of-words algorithm, a convolutional neural network, and feature combination, together with eight algorithms. As indicated in Table 9, the proposed prediction model achieved 73.68% data extraction with the use of LR_4 for Apple and 67.78% data extraction with the use of SVM_4 for Microsoft. In particular, the LR algorithm used for Apple achieved accuracy and F1-score of 0.7326 and 0.7360, respectively, which highlights its powerful capability in a two-class classification.

TABLE 9: Results for data extracted from Thomson Reuters.

| Algorithm	Company	Stock data (SD)	Bag of words (BOW)	Convolutional neural network	Feature combine model (FC)
		Accuracy/F1-score	Accuracy/F1-score	Accuracy/F1-score	Accuracy/F1-score
DT	Apple	0.4993/0.5079	0.4889/0.2460	0.6150/0.5814	0.6524/0.6486
	Microsoft	0.4868/0.1819	0.5065/0.5682	0.5817/0.5915	0.6587/0.6502
	Samsung	0.4707/0.4033	0.5654/0.5830	0.5529/0.6820	0.5680/0.6260
LR	Apple	0.5240/0.6727	0.5055/0.5083	0.6845/0.6911	0.7326/0.7368
	Microsoft	0.4757/0.1253	0.4430/0.4884	0.5274/0.5333	0.5467/0.5611
	Samsung	0.4947/0.3755	0.6054/0.6726	0.6541/0.6343	0.8053/0.7978
NB	Apple	0.5109/0.5123	0.4778/0.3896	0.5027/0.4497	0.5122/0.4030
	Microsoft	0.4840/0.0363	0.4557/0.4055	0.4599/0.2889	0.5000/0.5439
	Samsung	0.5293/0.5810	0.5144/0.6130	0.5048/0.6578	0.5117/0.6601
SGD	Apple	0.4672/0.1119	0.4778/0.4404	0.4974/0.6008	0.5366/0.2692
	Microsoft	0.5173/0.6819	0.4852/0.3222	0.6103/0.4231	0.6108/0.4193
	Samsung	0.5223/0.6541	0.4766/0.2000	0.5187/0.6751	0.5571/0.6548
SVM	Apple	0.5022/0.6121	0.5747/0.5542	0.6150/0.6129	0.6220/0.6173
	Microsoft	0.4784/0.1722	0.4439/0.4566	0.5232/0.5462	0.6875/0.6798
	Samsung	0.5008/0.3875	0.5047/0.5508	0.5135/0.6281	0.5352/0.6574
RF	Apple	0.5240/0.6554	0.5278/0.6083	0.5337/0.6232	0.6203/0.6243
	Microsoft	0.4840/0.2653	0.4177/0.1687	0.5325/0.2432	0.5721/0.5189
	Samsung	0.4872/0.3480	0.4485/0.1194	0.4766/0.3708	0.5052/0.5481
AB	Apple	0.5343/0.5384	0.4722/0.2963	0.4819/0.3333	0.6203/0.5420
	Microsoft	0.4687/0.2041	0.4768/0.0461	0.6147/0.6044	0.6202/0.6146
	Samsung	0.4857/0.3619	0.4579/0.2368	0.6010/0.6820	0.6338/0.6549
GB	Apple	0.5314/0.5129	0.4667/0,2941	0.6096/0.4748	0.6341/0.6591
	Microsoft	0.4743/0.2125	0.4728/0.1007	0.4557/0.4557	0.5385/0.2258
	Samsung	0.4872/0.4171	0.5192/0.6599	0.5337/0.6572	0.5622/0.6335
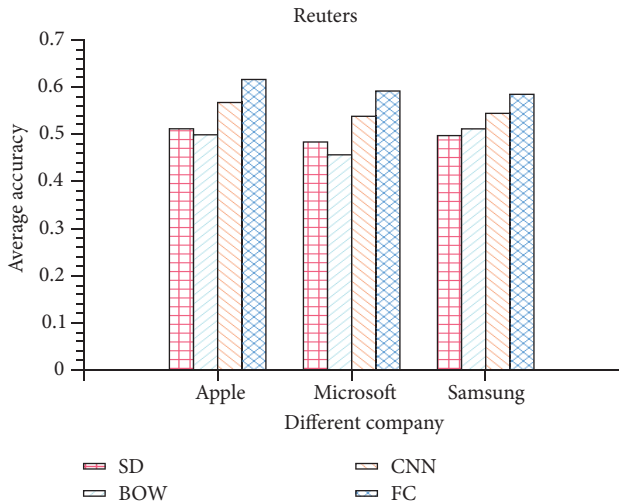
FIGURE 6: Average accuracy of prediction based on different feature sets for each company in Reuters.

To assess the effectiveness of our research, we compared the feature combination model that uses event tuples with the stock data approach and the bag-of-words algorithm. The results showed that the average accuracy of the event tuple-based model is better than that of the stock data approach and the bag-of-words algorithm by 10.01% and 10.87%, respectively. We also compared feature extraction using the stock data approach, the bag-of-words algorithm, and machine learning. The use of deep learning improved prediction accuracy by 5.25% and 6.11% over the levels achieved with stock data and bag of words, respectively. Therefore, we can conclude that using the proposed feature combination model and deep learning in feature extraction helps improve the accuracy with which stock price movement is predicted. These results also fully prove the effectiveness of embedding layer in feature extraction accuracy improvement.

4.2.2. CNN Case Study. CNN is an American basic cable and satellite television news channel owned by the Turner Broadcasting System. Three renowned companies on which these website reports were selected for the analysis. The first is Boeing, which is an American multinational corporation that designs, manufactures, and sells airplanes, rotorcraft, rockets, and satellites worldwide. The second company is Google, which is an American multinational technology

Table 10: Results for data extracted from CNN.

Algorithm	Company	Stock data (SD)	Bag of words (BOW)	Convolutional neural network	Feature combination model (FC)
		Accuracy/F1-score	Accuracy/F1-score	Accuracy/F1-score	Accuracy/F1-score
DT	Boeing	0.4950/0.3489	0.5413/0.6258	0.5414/0.6347	0.5591/0.5821
	Google	0.4858/0.5600	0.5182/0.3575	0.5692/0.4631	0.5491/0.5471
	Walmart	0.4996/0.4422	0.5300/0.4891	0.5088/0.4815	0.5722/0.5829
LR	Boeing	0.4923/0.2378	0.5113/0.4348	0.6923/0.6923	0.7596/0.5733
	Google	0.5041/0.6105	0.6235/0.6491	0.6250/0.6111	0.6563/0.6516
	Walmart	0.5290/0.5468	0.5412/0.5083	0.7320/0.7347	0.7719/0.7720
NB	Boeing	0.5162/0.6503	0.5489/0.5161	0.5197/0.6738	0.5564/0.7619
	Google	0.5112/0.6527	0.4899/0.5191	0.5587/0.5281	0.5344/0.4843
	Walmart	0.5044/0.3284	0.4794/0.5302	0.5206/0.5674	0.5300/0.6083
SGD	Boeing	0.4897/0.2407	0.5564/0.4870	0.5433/0.4727	0.5714/0.7220
	Google	0.5165/0.6769	0.5425/0.4375	0.5789/0.6364	0.5910/0.6327
	Walmart	0.5004/0.3804	0.5206/0.5280	0.5250/0.5581	0.5670/0.5922
SVM	Boeing	0.4897/0.3068	0.5714/0.6275	0.5385/0.6418	0.5940/0.7033
	Google	0.5077/0.6297	0.6032/0.6230	0.5749/0.5532	0.5789/0.5840
	Walmart	0.5118/0.4333	0.5000/0.5314	0.5200/0.5000	0.5614/0.5455
RF	Boeing	0.4906/0.1345	0.5113/0.3925	0.5096/0.5405	0.5276/0.5714
	Google	0.4923/0.1777	0.6032/0.6260	0.5223/0.5564	0.5951/0.5763
	Walmart	0.5118/0.5373	0.5300/0.3188	0.5300/0.4405	0.5000/0.5446
AB	Boeing	0.4945/0.3404	0.5197/0.5120	0.4961/0.5294	0.5673/0.5872
	Google	0.4876/0.3141	0.5101/0.6667	0.5893/0.5619	0.5951/0.6850
	Walmart	0.5073/0.5029	0.5150/0.5359	0.5497/0.5650	0.5619/0.5503
GB	Boeing	0.4899/0.3169	0.5118/0.4364	0.4409/0.4580	0.5000/0.5094
	Google	0.4912/0.3387	0.5870/0.5854	0.5625/0.5243	0.6032/0.6202
	Walmart	0.5003/0.4955	0.5050/0.5123	0.5361/0.5109	0.5497/0.5650

company that specializes in Internet-related services and products. The third is Walmart, an American multinational retail corporation that operates a chain of hypermarkets, discount department stores, and grocery stores.

The results on average accuracy are shown in Table 10 and Figure 7. These findings are very similar to those derived in the previous case study, and their comparison confirmed that the proposed feature combination model can outperform other feature selection models in stock price movement prediction. With regard to feature selection underlain by the bag-of-words algorithm, the CNN case study generated more robust results than did the Reuters case study. The average accuracy of the bag-of-words algorithm was lower than that of the stock data approach in the Reuters case study. In the CNN case study, the average accuracy levels of the proposed feature combination model were 57.94%, 58.79%, and 57.67% for Boeing, Google, and Walmart, respectively. In the case study on Reuters, an average accuracy that exceeds 60% was not achieved for any of the companies, illustrating that differences in data source directly affect stock price movement.

We chose a combination of deep learning and knowledge graph to build our feature selection model because this
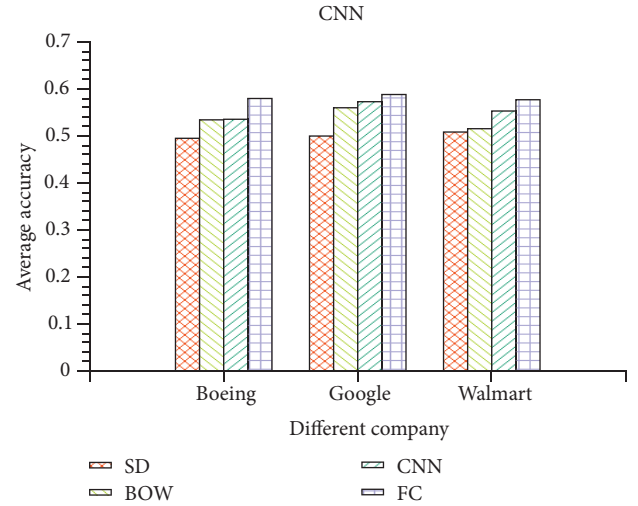

Figure 7: Average accuracy achieved on the basis of news from CNN and different feature set.

combination exhibited superior prediction performance in the comparison experiments involving other feature selection

strategies. The forecast results based on different features are shown in Table 10. We found that deep learning features are better than the use of bag-of-words and stock data. The correlation of event tuples with the stock market is relatively high in [5, 31], but the bag-of-words algorithm was relatively fragmented, and the correlation between event tuples and the stock market data was relatively weak. Combining event tuple features with deep learning significantly improves forecast results, indicating a close relationship between stock market movements and knowledge graph.

5. Discussion

5.1. Research Findings. We attempted to compare our prediction results with those made in previous studies wherein prediction was also based on the financial news of the renowned companies. However, the findings are incomparable because different studies use different datasets or algorithms, and these methods are difficult to investigate using financial news headlines. In the current research, we refrained from evaluating performance on one company dataset, this decision that enabled our knowledge graph method to exceed baseline predictions based only on the price by up to 3.6%. This also proves that feature extraction in deep learning is more effective than traditional machines. The combination of deep learning and knowledge graph fully integrates the semantic information in financial news, which effectively predicts the stock price movement of the renowned company.

This work demonstrated the application of deep learning and knowledge graph in finance. To the best of our knowledge, knowledge graph has been rarely applied in stock prediction. Because minimal financial training sets in knowledge graph are available, financial knowledge extraction is the main task for the organizational construction of a knowledge graph. Such extraction is critical for the understanding and processing of deep semantics in the event tuple, which also directly influences the feature extraction of financial news.

5.2. Implications for Business Activities. The utility of a feature combination model based on knowledge graph is not limited to financial analysis. Currently, knowledge graph data are available for medical diagnosis, speech recognition, precise marketing, and financial risk control [4]. Our model could also be applied in these areas.

Antifraud activities are an important part of finance. Applying our model on the basis of knowledge graph on customers helps organize all relevant knowledge fragments through deep semantic analysis and reasoning, which can be verified with a customer's bank information. Moreover, customers typically use keywords to search for products, and knowledge graph can provide relevant information to a customer. If a complete knowledge system of users is described and collected, a system would better understand and analyze user behaviors.

5.3. Limitations and Future Work. A deep learning model uses supervised learning and needs a dataset with sufficient labels, but our datasets do not work well with deep learning as a small number of financial news articles match stock

data. At present, knowledge graph embedding inevitably produces a loss in news semantics, either because of a learning principle or the tokenization of knowledge representation for vectorization. Thus, continuous representation in a knowledge graph is still a huge challenge. We attempted to apply regression using the above-mentioned algorithms, but the results indicated poor performance. More specifically, the results for Apple validated our expectations.

The application of large-scale knowledge graph is still relatively limited, and knowledge graph for smart search, smart question-and-answer tasks, social media, and other areas are in their initial stages, with considerable room for improvement. The following advantages of knowledge graph should be considered: (a) the effective organization and expression of semistructured data, (b) knowledge reasoning, and (c) the expansion of cognitive ability for incorporating deep learning. Traditional knowledge elements (entities, relationship, attributes), extraction technologies, and methods have achieved good results in limited areas, but because of numerous constraints and poor scalability, knowledge graph is not fully functional for financial forecasting because of numerous constraints and poor scalability.

6. Conclusions

Stock movement prediction of the renowned company is a daunting task because stock prices are affected by many factors. This research presented a novel method for integrating knowledge graph embedding with stock market prediction. The contributions of this study can be summarized as follows. First, we developed a novel feature combination model that constructs a feature mapping vector for each tuple–news pair by simultaneously considering the diversity of entities and relationships. Second, the feature combination model was successfully applied to different types of companies and datasets and exhibited good performance in classification tasks.

Stock market prediction of the renowned company grounded in knowledge graph is an interesting topic for business activities. Given that the study of knowledge graph for feature set is still in its infancy, we expect it to be applied in a wide range of academic research. More companies will also earn profits and create more opportunities through the use of knowledge graph in the feature set.

Data Availability

The data used in this study can be accessed via https://github.com/linechany/knowledge-graph.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] B. G. Malkiel and E. F. Fama, "Efficient capital markets: a review of theory and empirical work," *The Journal of Finance*, vol. 25, no. 2, pp. 383–417, 1970.

[2] D. Galai and R. W. Masulis, "The option pricing model and the risk factor of stock," *Journal of Financial Economics*, vol. 3, no. 1-2, pp. 53–81, 1976.

[3] Q. Wang, Z. Mao, B. Wang, and L. Guo, "Knowledge graph embedding: A survey of approaches and applications," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 12, pp. 2724–2743, 2017.

[4] H. Paulheim and P. Cimiano, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 8, no. 3, pp. 489–508, 2016.

[5] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," in *Proceedings of the 24th International Joint Conference on Artificial Intelligence, IJCAI 2015*, pp. 2327–2333, Argentina, July 2015.

[6] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Knowledge-driven event embedding for stock prediction," in *Proceedings of the 26th International Conference on Computational Linguistics, COLING 2016*, pp. 2133–2142, Japan, December 2016.

[7] A. Bordes, N. Usunier, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in NIPS*, vol. 26, pp. 2787–2795, 2013.

[8] H. Lin, Y. Liu, W. Wang, Y. Yue, and Z. Lin, "Learning entity and relation embeddings for knowledge resolution," *Procedia Computer Science*, vol. 108, pp. 345–354, 2017.

[9] G. Ji, K. Liu, S. He, and J. Zhao, "Knowledge graph completion with adaptive sparse transfer matrix," in *Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI 2016*, pp. 985–991, USA, February 2016.

[10] G. Ji, S. He, L. Xu, K. Liu, and J. Zhao, "Knowledge graph embedding via dynamic mapping matrix," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pp. 687–696, Beijing, China, July 2015.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[12] Q. Li, Y. Chen, J. Wang, and H. Chen, "Web media and stock markets: a survey and future directions from a big data perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 4347, p. 1, 2017.

[13] B. Wuthrich, V. Cho, S. Leung, D. Permunetilleke, K. Sankaran, and J. Zhang, "Daily stock market forecast from textual Web data," in *Proceedings of the 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.98CH36218), SMC'98*, vol. 3, pp. 2720–2725, October 1998.

[14] G. Pui Cheong Fung, J. Xu Yu, and W. Lam, "Stock prediction: integrating text mining approach using real-time news," in *Proceedings of the 2003 IEEE International Conference on Computational Intelligence for Financial Engineering, CIFEr 2003*, vol. 2003, pp. 395–402, China, March 2003.

[15] W. Antweiler and M. Z. Frank, "Is all that talk just noise? the information content of internet stock message boards," *SSRN Electronic Journal*, 2001.

[16] M.-A. Mittermayer and G. F. Knolmayer, "NewsCATS: A news categorization and trading system," in *Proceedings of the 6th International Conference on Data Mining, ICDM 2006*, pp. 1002–1007, China, December 2006.

[17] R. P. Schumaker and H. Chen, "Textual analysis of stock market prediction using breaking financial news: the AZFin text system," *ACM Transactions on Information and Systems*, pp. 1–29, 2009.

[18] S. S. Groth and J. Muntermann, "An intraday market risk management approach based on textual analysis," *Decision Support Systems*, vol. 50, no. 4, pp. 680–691, 2011.

[19] T. Vu, S. Chang, Q. Ha, and N. Collier, "An experiment in integrating sentiment features for tech stock prediction in twitter," in *Proceedings of the Workshop on Information Extraction And Entity Analytics on Social Media Data*, vol. 3, pp. 23–38, 2012.

[20] B. Wang, H. Huang, and X. Wang, "A novel text mining approach to financial time series forecasting," *Neurocomputing*, vol. 83, pp. 136–145, 2012.

[21] Q. Li, T. Wang, P. Li, L. Liu, Q. Gong, and Y. Chen, "The effect of news and public mood on stock movements," *Information Sciences*, vol. 278, pp. 826–840, 2014.

[22] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for stock movement prediction," *Expert Systems with Applications*, vol. 42, no. 24, pp. 9603–9611, 2015.

[23] B. Weng, M. A. Ahmed, and F. M. Megahed, "Stock market one-day ahead movement prediction using disparate data sources," *Expert Systems with Applications*, vol. 79, pp. 153–163, 2017.

[24] J. Eberhard, J. F. Lavin, and A. Montecinos-Pearce, "A network-based dynamic analysis in an equity stock market," *Complexity*, vol. 2017, Article ID 3979836, 16 pages, 2017.

[25] Z. Zhang, Y. Zhang, D. Shen, and W. Zhang, "The dynamic cross-correlations between mass media news, new media news, and stock returns," *Complexity*, vol. 2018, 11 pages, 2018.

[26] P. C. Tetlock, "Giving content to investor sentiment: The role of media in the stock market," *Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, 2007.

[27] K. Chen, P. Luo, L. Liu, and W. Zhang, "News, search and stock co-movement: Investigating information diffusion in the financial market," *Electronic Commerce Research and Applications*, vol. 28, pp. 159–171, 2018.

[28] F. Z. Xing, E. Cambria, and Y. Zhang, "Sentiment-aware volatility forecasting," *Knowledge-Based Systems*, vol. 176, pp. 68–76, 2019.

[29] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, "Technical analysis and sentiment embeddings for market trend prediction," *Expert Systems with Applications*, vol. 135, pp. 60–70, 2019.

[30] A. Khadjeh Nassirtoussi, S. Aghabozorgi, T. Ying Wah, and D. C. L. Ngo, "Text mining for market prediction: a systematic review," *Expert Systems with Applications*, vol. 41, no. 16, pp. 7653–7670, 2014.

[31] M. Kraus and S. Feuerriegel, "Decision support from financial disclosures with deep neural networks and transfer learning," *Decision Support Systems*, vol. 104, pp. 38–48, 2017.

[32] H. S. Sim, H. I. Kim, and J. J. Ahn, "Is deep learning for image recognition applicable to stock market prediction?" *Complexity*, vol. 2019, 10 pages, 2019.

[33] M. Jasemi, A. M. Kimiagari, and A. Memariani, "A modern neural network model to do stock market timing on the basis of the ancient investment technique of Japanese candlestick," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3884–3890, 2011.

[34] A. Fader, S. Soderland, and O. Etzioni, "Identifying relations for open information extraction," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP 2011*, pp. 1535–1545, Association for Computational Linguistics, UK, July 2011.

[35] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," pp. 1–12, 2013, https://arxiv.org/abs/1301.3781.

[36] X. Lin, Z. Yang, and Y. Song, "Intelligent stock trading system based on improved technical analysis and Echo State Network," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11347–11354, 2011.

[37] K.-J. Kim and I. Han, "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index," *Expert Systems with Applications*, vol. 19, no. 2, pp. 125–132, 2000.

[38] Y. Yu, W. Duan, and Q. Cao, "The impact of social and conventional media on firm equity value: A sentiment analysis approach," *Decision Support Systems*, vol. 55, no. 4, pp. 919–926, 2013.

[39] M. Hagenau, M. Liebmann, and D. Neumann, "Automated news reading: Stock price prediction based on financial news using context-capturing features," *Decision Support Systems*, vol. 55, no. 3, pp. 685–697, 2013.

[40] A. Aizawa, "An information-theoretic perspective of tf-idf measures," *Information Processing & Management*, vol. 39, no. 1, pp. 45–65, 2003.

[41] A. Khadjeh Nassirtoussi, S. Aghabozorgi, T. Ying Wah, and D. C. Ngo, "Text mining of news-headlines for FOREX market prediction: a Multi-layer dimension reduction algorithm with semantics and sentiment," *Expert Systems with Applications*, vol. 42, no. 1, pp. 306–324, 2015.

[42] D. Peramunetilleke and R. K. Wong, "Currency exchange rate forecasting from news headlines," *Australian Computer Science Communications*, vol. 24, no. 2, pp. 131–139, 2002.

[43] Y. Kim, "Convolutional neural networks for sentence classification," 2014, https://arxiv.org/abs/1408.5882.

[44] Z. Wang, J. Zhang, J. Feng, and Z. Chen, "Knowledge graph and text jointly embedding," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014*, pp. 1591–1601, Qatar, October 2014.

[45] J. Xu, X. Qiu, K. Chen, and X. Huang, "Knowledge graph representation with jointly structural and textual encoding," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI 2017*, pp. 1318–1324, Australia, August 2017.

[46] R. Xie, Z. Liu, J. Jia, H. Luan, and M. Sun, "Representation learning of knowledge graphs with entity descriptions," in *Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI '16*, pp. 2659–2665, February 2016.

[47] F. Tian, B. Gao, E. Chen, and T. Liu, "Learning better word embedding by asymmetric low-rank projection of knowledge graph," *Journal of Computer Science and Technology*, vol. 31, no. 3, pp. 624–634, 2016.

[48] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting: a survey," *Artificial Intelligence Review*, vol. 50, no. 1, pp. 49–73, 2018.

[49] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and *F*-score, with implication for evaluation," in *Advances in Information Retrieval*, D. E. Losada and J. M. Fernández-Luna, Eds., vol. 3408 of *Lecture Notes in Computer Science*, pp. 345–359, Springer, Berlin, Germany, 2005.