

# ***Electronic word-of-mouth effects on studio performance leveraging attention-based model***

Yang Liu, Hao Fei, Qingguo Zeng, Bobo Li, Lili Ma, Donghong Ji, Joaquín Ordieres Meré

## **Abstract**

While existing studies have established the relationship between electronic word-of-mouth (eWOM) and studio performance, limited research has been conducted to demonstrate how the attention-based model applies to the motion picture industry. In this study, examining a review corpus of seven Hollywood studios, we proved that deep learning with the attention mechanism has the best accuracy in both eWOM and stock price movement. We present both a hierarchical two-layer attention network and hierarchical convoluted attention network (HCAN), which quantify the importance of crucial eWOM features in capturing valuable information from audience members' reviews. Further, comparing the two case studies, we determined that the HCAN model is superior to both machine learning and attention-based models. Our work helps to highlight the business value of the attention-based model and has implications for studio business decisions.

**Keywords** Electronic word-of-mouth · Audience review · Stock market · Deep learning · Attention mechanism

## **1 Introduction**

As the development of artificial intelligence (AI) technology applications accelerates, researchers have a great interest in natural language processing (NLP) and text mining in order to accurately extract important information from online reviews [1], because online reviews on social media are considered a valuable source of information for most firms [2]. Online review often sought out by many potential consumers who understand the experience of other consumers to make better purchasing decisions [3]. Customers can share their feedback on product experiences on the social media platform, namely, electronic word of mouth (eWOM) [4]. Gretzel and Yoo [5] also proved that eWOM is an important indicator for evaluating online reviews. Consequently, firms are increasingly making business decisions using the eWOM feature in online reviews [6].

This is true especially in the motion picture industry. Audiences frequently interact through Internet Movie Database (IMDb), where they can share their feedback on a movie experience. Systematic research on review impact has provided important decision support for studios. Previous research [7, 8] has fully examined the positive relationship between online word-of-mouth and movie box office performance, and existing literature [9, 10] has demonstrated the positive relationship between eWOM expressing customers' sentiments and studio stock market performance, which proved that eWOM is still a powerful force for the studios. In general, positive eWOM has increased movie box office numbers and stock market performance of studios; negative eWOM has the opposite result. Hence, we examine eWOM effects on both eWOM and the stock market, using two case studies in seven studios.

Currently, audience sentiment is the foremost factor of eWOM about the studios [7, 8]. Many researchers realize that the traditional machine learning model does not provide sufficient decision support for studios [11]. Most of the work uses latent Dirichlet allocation [12] and non-negative matrix factorization [13] to extract the eWOM feature from audience reviews. These methods involve processing text based on a bag-of-words approach, which calculates the frequency of words as  $n$ -grams and ignores context-associated semantic information in audience reviews. To address this problem, deep learning provides powerful learning machinery that attracts the application of natural language processing in audience reviews analysis [14].

Since most audience sentiment is implicit [15], sentiment analysis of eWOM using the deep learning model shows an improvement over the traditional machine learning method, which only pays attention to the content of the text comments and ignores the crucial sentiment characteristic in audience reviews. Hence, many researchers have been interested in the attention mechanism in deep learning [16], which has been widely used, such as machine translation, question answering, multimedia description, and image recognition [17]. This method is first utilized in machine translation that assigns different weights to each word in a sentence. Meanwhile, previous works [18, 19] show the following limitations: (1) the eWOM characteristics on audience reviews are only reflected on the word level or sentence level; (2) each audience review consists of multiple characters, word and sentences. These works did not consider these the fine-grain eWOM characteristics in the audience review. In this work, we adopted the attention mechanism to encode the crucial eWOM characteristics in the audience review, which can address the above limitations.

Overall, our work furnishes the following contributions. First, we establish a hierarchical two-layer attention network (HTAN) that generates character-level and word-level representation in the two case studies. Specifically, audience reviews are feed into the bidirectional long short-term memory (bi-LSTM) to extract eWOM features which adopts a bag-of-words approach and word embedding. Hereafter, the word vectors are input into the softmax layer to gain the final classification label. Second, we develop a hierarchical convoluted attention network (HCAN) model in two case studies. We successfully gain the best performance through deep learning with the attention mechanism, which can emphasize the importance of embedding important words in each review and enriches the eWOM feature extraction in the end-to-end. Finally, the HCAN model promises to strengthen the power in two cases between eWOM and the stock market. Our proposed method achieves better performance than that exhibited by

the attention-based models that using attention-based convolutional neural network (A-CNN) [20], attention-based long short-term memory neural network (A-LSTM) [21], user and product attention neural network (UPA) [18] and hierarchical user attention and product attention neural network (HUAPA) [22], which also highlights the superiority of the attention mechanism. The additional analysis is used to examine the effectiveness of the method. In summary, this work can provide an applicable framework for studio performance, and our framework also should benefit academics and industry practitioners who provide insightful decision support for the motion picture industry.

The remainder of the paper is organized as follows. Section 2 is a review of the existing literature. Data and material are shown in Sect. 3. Section 4 describes our proposed method. Afterword, we expound the experiment results of two case studies in Sect. 5. Finally, Sect. 6 discusses the implication for business research, limitation and future work, while Sect. 7 concludes.

## 2 Literature review

### 2.1 Linking eWOM with studio performance evaluation

Table 1 reviews publications that highlight eWOM performance effects. Prior to the movies was released, the firms utilized a survey of eWOM, which detects the sentiment performance of the audience, to reduce economic risks. Xun and Guo [31] explored whether positive and negative aspects of eWOM would impact target firms. The sensitivity ratio of eWOM is bigger, and users are more optimistic about the firms: they expect that the firms have higher returns and less volatility. However, the above works do not include any variables that capture the stock market of the studios, and our work can fill this gap in studios performance.

Scholars measured the impact of eWOM through the parameters [11]. They also recognized the audience sentiment aspect of eWOM that may have a greater impact on the customer purchase decision. Audience sentiment is relatively easy to understand and can be either negative or positive, which becomes the metric in eWOM [32]. Kim et al. [33] examined the eWOM effects through consumer product reviews, which demonstrated the positive relationship between customer sentiment and attitudes toward the firm's performance. Moreover, Yin et al. [34] examined the relationship between sentiment and consumer, which proved that positive sentiment facilitates consumer product purchases. Previous works [10, 35] further illustrated that consumer sentiment directly affected a company's stock returns. We specifically investigated audience sentiment in

**Table 1** eWOM performance effects, the motion picture industry literature review, and the contributions of our research

Studies	Theoretical foundations	Method	eWOM measure	Case measure
[23]	Market motivation	Regression	Average rating, number of ratings	Box office revenue
[8]	Market signal	Regression	Number of posts, reviewer valence	Box office revenue
[7]	Retail feedback signal	Dynamic simultaneous equation system	Retail sales	Box office sales
[24]	Market sentiment	Regression	Reviewer valence	Review helpfulness
[3]	Market dialectics	Two sided argument	Reviewer's arguments, number of ratings	Review helpfulness
[25]	Market motivation	Regression	Average rating, reviewer valence	Box office revenue
[26]	Market sentiment	Regression	Review valence, review quantity	Review helpfulness
[27]	Market motivation	Machine learning	Average rating, reviewer valence	Box office revenue
[28]	Market motivation	Machine learning	Average rating, reviewer valence	Box office sales
[29]	Market sentiment	Regression, machine learning	Reviewer volume, reviewer valence	Box office sales
[30]	Roger's innovation diffusion	Regression	Type of review, review volume	Box office revenue
The present study	Sentiment finance interface	Machine learning, deep learning	Average rating, reviewer valence, review quantity	eWOM, stock price movement

the valence feature of eWOM, which fully taps the needs of the audience to determine precise marketing for the studios. Then, we examined the audience sentiment of eWOM, which influences studio performance. We propose a way to go beyond the valence of eWOM and instead utilize the attention mechanism to improve the effect of eWOM on studios.

## 2.2 eWOM and stock price movement in studio performance

In behavioral finance, stock market of companies may be influenced by customer sentiment on social media [36]. To analyze customers' word of mouth, researchers use various types of sentiment analysis techniques to improve the performance of predictive stock market [37]. Generally, positive word-of-mouth drives to increase in stock price. On the contrary, negative word-of-mouth drives to decrease in stock price. Tetlock [38] found that public sentiment can predict stock market of the company. Li et al. [39] proposed quantitative trading strategies through public sentiment orientation. Nguyen et al. [40] constructed the sentiment topic model from financial news, and the experimental results showed that the forecast performance had improved significantly by 3.03%. The above literature fully illustrates the positive effect of emotional symbol stock prediction. However, existing literature did not consider special companies (e.g., studio). We thereby utilize the eWOM feature of audiences to predict the stock market of the studios.

Many sentiment words directly reflect the eWOM feature. Li et al. [41] captured sentiment words from the financial news, which reveals the impact of public sentiment on stock trading. However, this work ignores the impact of the recessive sentiment words on the stock. If eWOM feature is implicit rather than direct sentiment discourses, using sentiment analysis techniques to predict the stock market is relatively limited, which requires us to consider better technical support. Si et al. [42] constructed a semantic network from tweets to calculate the public sentiment for predicting stock market. Due to the noise contained in the audience review, it is difficult to accurately capture the eWOM feature for stock market prediction. Previous literature [43] had proven that the attention-based model can better capture implicit sentiment words in complex text. Hence, we can contract the predictive framework through the attention-based model.

## 2.3 Machine learning for studio performance

Audience reviews have been dominated by traditional machine learning, which is widely used for classification in online review analysis [44]. Traditional machine learning methods mainly depend on feature extraction in human subjective design, which primarily contributes to the best performance for classifiers or regression models. In previous studies [45], machine learning-based algorithms have been developed to automatically extract useful information related to studio performance. A novel approach to motion picture box office forecasting using social network service



data and machine learning-based algorithms was proposed by [46]; the experiment result made noticeable improvements to the forecasting accuracies of all three models. Du et al. [47] predicted box office revenue from microblogs through machine learning, which showed that the prediction result is more accurate.

Currently, deep learning provides a powerful learning capability based on machine learning [14]. There are two major kinds of neural network architectures in deep learning: convolutional neural network and recurrent neural network. Convolutional neural network (CNN) is a specialized structure of multi-layer perceptron (MLP); it is proficient in extracting local patterns in the corpus. This network is suitable for distinguishing indicative phrases or idioms of fixed length in the corpus. Recurrent neural network (RNN) is a specialized structure for sequential data in the corpus. There are two variants in RNN, LSTM and gated recurrent units (GRU) [14]. These variants take a sequence of words or sentences as input and generate a fixed size vector in the corpus.

The deep learning model had also been applied to stock price forecasts for firms' performance [48]. Accuracy using deep learning is improved by nearly 6% on Standard & Poor's 500 index compared to previous state-of-the-art baseline methods [49]. Kraus and Feuerriegel [50] utilized deep neural networks and transfer learning for financial decision support; the results indicated a better performance compared to traditional machine learning. Evidently, the deep learning model performs well in feature learning related to various prediction tasks. Our work, therefore, applies state-of-the-art attention mechanism in deep learning for the studios.

## 2.4 Attention mechanism for studio performance

Although deep learning has the ability to learn powerful text representation through multiple hidden layers, it cannot effectively capture the characteristics of sentiment words when the target sentiment words are far away from the target. Attention mechanism can be seen as a way to make the neural network learn better by letting the network know the specific location and weight of the word vector. It assigns attention scores to context words by modeling between target and context. Yin et al. [20] utilized the attention mechanism that integrates mutual influence between sentences into CNN, and the results achieve state-of-the-art performance on answer selection, paraphrase identification and textual entailment. Wang et al. [21] accurately captured the important sentiment characteristics of a given target sentence based on the attentional mechanism. Chen et al. [18] proposed an attention-based sentiment classification model, which effectively captures

global user preferences and product characteristics. Wu et al. [22] proposed HUAPA to generate two representations to encode user and product information; experimental results show that their model is significantly better than other recent methods on the both IMDb and Yelp datasets. He et al. [51] utilized an attention-based LSTM model to combine the crucial information of sentences, obtained from document-level sentiment classification, to achieve improved performance of sentiment classification. Liu and Zhang [52] combined left and right context using the two LSTM attention models, which induced an attention value of the whole sentence.

The above literature primarily relates to sentiment analysis; however, use of the attention mechanism can also extend to the stock market. Hu et al. [53] designed a hybrid attention network that predicted stock trends according to the chronological order of the news; this framework can significantly improve annualized returns. Liu et al. [54] adopted a two-level attention mechanism to quantify the importance of words and sentences in given financial news. This is consistent with our proposed model, which ignores the importance of character levels and cannot understand the fine-grained sentiment characteristic. Furthermore, previous literature [55, 56] also demonstrated the importance of character-level attention in NLP tasks. In summary, our work employed attention mechanisms for both eWOM and stock market based on HAN [19]. We combined the characteristics of our tasks and datasets to deeply transform the HAN model. The first attention layer in the HTAN model encodes character level, which fully captures the important characters in the words, and its second attention layer encodes the word level in the whole words, which also captures list the important words in the audience reviews. Meanwhile, the HCAN model can provide the best results, due to replacing RNN with CNN based on the HAN model. An ablation experiment was performed on the model to prove the effectiveness of the attention.

## 3 Dataset and material

### 3.1 Audience review

Audience reviews for this work were collected from IMDb (<https://www.imdb.com/>), which provides a large amount of data related to movie information. IMDb is the most well-known online review site. We designed a web crawler by Python with Scrapy, which captures these reviews, a total 60,012 reviews from 95 movies between 4 November 2013 and 26 July 2018. Our movie volume was commonly used in previous works [30, 57]. Figure 1 shows an example of an audience member's review on IMDb. It also consists of the review published at specific time intervals,

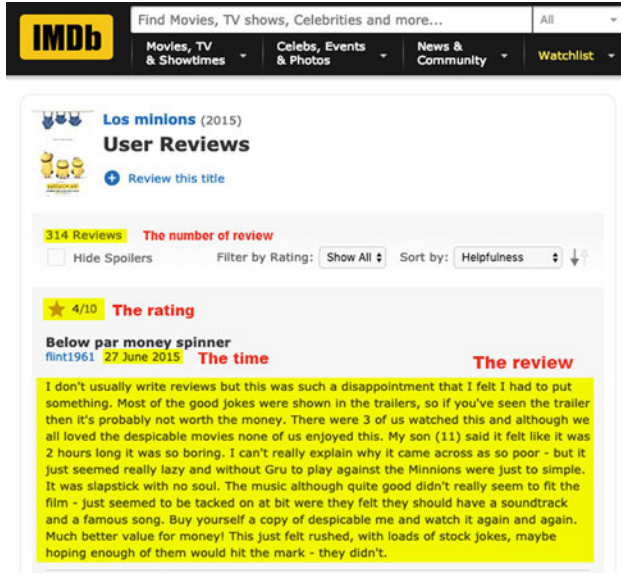


Fig. 1 Example of audience member's review on IMDb

with each review report accompanied by a title, a release date, and audience rating.

### 3.2 Stock data

The list of studio names is shown in Table 2. Historical stock price data for the studios are retrieved from Yahoo Finance; each studio stock consists of open price, high price, low price, close price, and adjusted close price for each transaction date.

### 3.3 Data pre-processing

We gathered audience reviews of movies from the same studio from IMDb. For instance, *A Cure for Wellness*, *Goodbye Christopher Robin*, and *The Mountain Between Us* were produced by Fox; audience reviews were combined in the same corpus. The dataset contains audience reviews, audience ratings, and review content. We tagged audience reviews through the rating-based method. An audience rating less than 5 means that the sentiment of this

Table 2 Quotes and studios' names

Studios	Studios name	Stock
Disney	The Walt Disney Corporation	DIS
Fox	Twentieth Century Fox Film Corporation	FOX
Lionsgate	Lions Gate Entertainment Corporation	LGF.A
Paramount	Paramount Pictures Corporation	PGRE
Sony/Columbia	Sony/Columbia Pictures	SNE
Universal	Universal Studios	UVV
Warner	Warner Bros. Entertainment, Inc.	TWX

review is negative; then the label given is 0. In contrast, if the audience rating is equal to or greater than 5, the sentiment is positive, and the label given is 1.

Table 3 shows matching statistics between audience reviews and stock prices for each studio. The release date of review is used to align the corresponding trading day from a temporal series. We removed non-trading days, such as weekends and holidays; unpaired stock data and audience reviews were also removed, since the stock market only occurs on trading days (except weekends and holidays). If the release date or calculation date happened on a non-trading day, we pushed the release date back 1 day or 2 days.

There were fewer reviews associated with the stock market than associated with eWOM. The target output consists of a binary variable, where a value of 1 indicates that the closing price that 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  $t$ , which is defined in the following equation:

$$\hat{y} = \begin{cases} 0 & p > 0 \\ 1 & \text{else} \end{cases}, \quad (1)$$

where  $\hat{y}$  represents the predicted label in each candidate pair, and  $p$  denotes distribution probability. Afterward, according on the principle of data preparation for time series, we split 80% of the samples for training and the remaining 20% (happening later than dataset used for training), for testing purposes, as presented in Table 4. Meanwhile, we also calculated the time interval of review corpus in Table 5.

## 4 Methodology

This section introduces our methodology in the deep learning method. We first summarized our baseline from traditional machine learning and deep learning. Specifically, we compared traditional machine learning techniques using a bag-of-words approach and deep learning

Table 3 Summary statistics between audience reviews and stock price in one trading day

Studios	Min	Mean	Max	SD	Trading days
Disney	1	8.18	415	20.86	1134
Fox	1	10.04	334	22.34	786
Lionsgate	1	2.92	47	3.94	598
Paramount	1	4.07	54	4.43	780
Sony/Columbia	1	3.93	51	5.56	610
Universal	1	5.64	100	7.88	760
Warner	1	8.23	178	13.66	961

**Table 4** Data information of review corpus in two cases of studios' performance

	eWOM			Stock price movement		
	Training	Testing	Total	Training	Testing	Total
Disney	12,124	3031	15,155	7420	1855	9275
Fox	9958	2490	12,448	6316	1579	7895
Lionsgate	2194	548	2741	1399	350	1749
Paramount	4879	1220	6099	2540	635	3175
Sony/Columbia	2959	740	3699	1918	480	2398
Universal	5810	1453	7263	3430	858	4288
Warner	10,085	2521	12,606	6326	1581	7907

**Table 5** Time interval of review corpus in two cases of studios' performance

	Time interval		eWOM				Stock price movement			
			Training		Testing		Training		Testing	
Disney	18.4.2013	17.4.2018	18.4.2013	8.1.2018	8.1.2018	17.4.2018	18.4.2013	22.12.2017	22.12.2017	17.4.2018
Fox	9.5.2014	24.7.2018	9.5.2014	25.12.2017	25.12.2017	24.7.2018	9.5.2014	20.12.2017	20.12.2017	24.7.2018
Lionsgate	11.11.2013	25.7.2018	11.11.2013	25.7.2018	15.9.2017	25.7.2018	11.11.2013	15.9.2017	15.9.2017	25.7.2018
Paramount	29.10.2014	26.7.2018	29.10.2014	23.8.2017	23.8.2017	26.7.2018	29.10.2014	3.10.2017	3.10.2017	26.7.2018
Sony/Columbia	11.4.2014	25.7.2018	11.4.2014	21.12.2017	21.12.2017	25.7.2018	11.4.2014	18.12.2017	19.12.2017	25.7.2018
Universal	13.5.2013	25.7.2018	13.5.2013	21.8.2017	22.8.2017	25.7.2018	13.5.2013	8.9.2017	8.9.2017	25.7.2018
Warner	11.6.2013	14.7.2018	11.6.2013	29.10.2017	29.10.2017	26.7.2018	11.6.2013	26.10.2017	26.10.2017	14.6.2018

using word embedding. While we propose attention mechanism in Sects. 4.2.3 and 4.2.4, it is a process that enables an inductive method of classification from the related task. Figure 2 illustrates this pipeline to providing decision support using eWOM features for studios. We evaluate the performance of the model through accuracy, sensitivity, and specificity on the confusion matrix. These results help marketers and managers in studios make better business decisions.

#### 4.1 Benchmark from traditional machine learning

This section briefly introduces the traditional machine learning model with bag-of-words in feature extraction. Figure 3 details this process, and we calculate the frequency of occurrences of terms in the document term matrix, in order to reduce dimension for removing sparse and put in the frequency in the matrix. This matrix represents the feature for weighting the actual value by frequency-inverse document frequency (TF-IDF) [58]; it is a numerical statistic that intends to reflect how important a word is to a review corpus. TF-IDF is often used as a weighting factor in searches of information retrieval, text mining, and user modeling [58]. We create different machine learning algorithms for two cases in the studios' performance. In this evolution task, we utilize traditional

machine learning as a baseline, which is consistent with previous works [50, 59].

### 4.2 Deep learning

#### 4.2.1 LSTM unit

LSTM is a special form of recurrent neural network; it is usually used to process sequence data and avoids the gradient explosion of traditional recurrent neural network or the problem of gradient disappearance [60]. LSTM captures long-distance dependencies in a sequence through the mechanism for memory cells and gates. The calculation of LSTM is as follows:

$$i_t = \sigma(W_i x_t + U_i h_{t-1}), \quad (2)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1}), \quad (3)$$

$$o_t = (W_o x_t + U_o h_{t-1}) \quad (4)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1}), \quad (5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t, \quad (6)$$

$$y_t = \tanh(o_t \odot c_t), \quad (7)$$

where  $\sigma$  represents the logistic sigmoid function;  $\odot$  represents the operator of the point multiplication;  $i_t, f_t, o_t$ , and  $c_t$  mean the input gate and forget gate, respectively. The output gate is at time  $t$ , and activation vectors of the



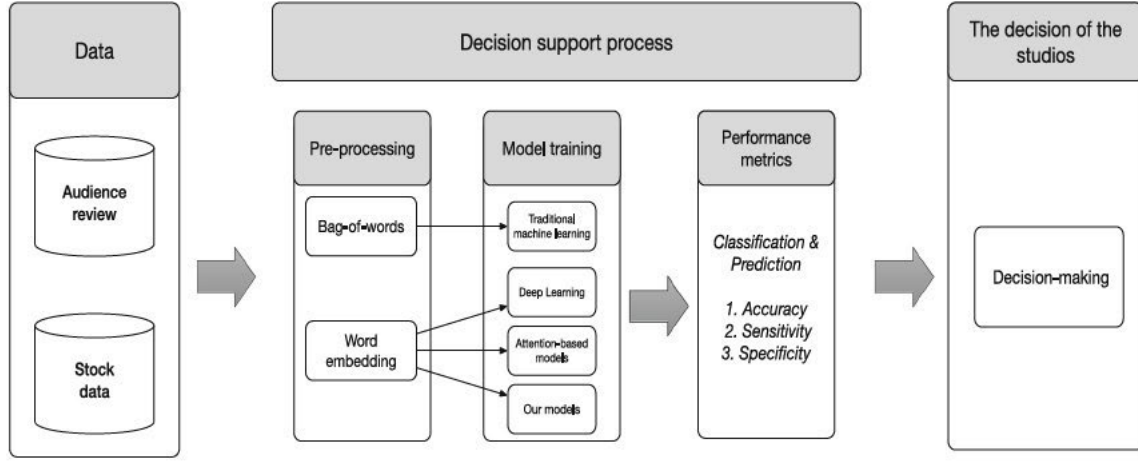


Fig. 2 Illustrative pipeline for the decision making of studios from narrative materials

memory unit are the same dimension as the hidden layer vector  $h_t$ . The weight matrices  $W$  and  $U$  are the model parameters in LSTM.

#### 4.2.2 GRU unit

GRU [60] has excellent performance in sequence modeling. This algorithm is a variant of RNN that utilizes the gate mechanism to record the current state of the sequence. It can solve problems such as gradient explosion. The previous state  $h_{t-1}$  and the current new state  $\tilde{h}_t$  information is computed by linear interpolation. Thus, the new state  $h_t$  is computed by GRU as:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t, \quad (8)$$

The update gate  $z_t$  is used to determine how much past information is retained and how much new information is added:

$$z_t = \sigma(W_z d_t + U_z h_{t-1} + b_z), \quad (9)$$

The  $t$  and the  $d_t$  represent input word vectors, so the current update gate  $\tilde{h}_t$  is represented as:

$$\tilde{h}_t = \tanh(W_h d_t + r_t * (U_h h_{t-1}) + b_h), \quad (10)$$

The reset gate  $r_t$  decides how much past information is applied to the candidate state. If  $r_t$  is 0, hence, the gate forgets all previous states. The  $r_t$  is represented as:

$$r_t = \sigma(W_r d_t + U_r h_{t-1} + b_r). \quad (11)$$

#### 4.2.3 Hierarchical two-layer attention network (HTAN)

The hierarchical two-layer attention network model is a hierarchical network of attention, which consists of the following main components: character level, word level,

and classification model. Figure 4 shows the architecture of a hierarchical two-layer attention network. This model is inspired by (HAN) [19].

##### 1. Character level

*Character encoder* Each word contains  $T_i$  characters, given a word with characters  $c_{it}, t \in [0, T]$ . We encode the character to vectors by embedding matrix  $W_c$ ,  $x_{it} = W_c c_{ij}$ . For instance, for the word for “Happy,” it enters characters “H” “a” “p” “p” “y” in turn. Then we encode the information of  $c_{it}$  by a bidirectional LSTM; thus, we can get its hidden layer representation. The specific calculation method is as follows:

$$x_{it} = W_c c_{it}, \quad t \in [1, T], \quad (12)$$

$$\vec{n}_{it} = \overrightarrow{\text{LSTM}}(x_{it}), \quad t \in [1, T], \quad (13)$$

$$\overleftarrow{n}_{it} = \overleftarrow{\text{LSTM}}(x_{it}), \quad t \in [T, 1], \quad (14)$$

$$n_{it} = [\vec{n}_{it} : \overleftarrow{n}_{it}], \quad (15)$$

where  $\overrightarrow{\text{LSTM}}(\cdot)$  and  $\overleftarrow{\text{LSTM}}(\cdot)$  denote the forward LSTM and the backward LSTM, respectively.  $[\vec{n}_{it} : \overleftarrow{n}_{it}]$  denotes a concatenating operation on  $\vec{n}_{it}$  and  $\overleftarrow{n}_{it}$ .

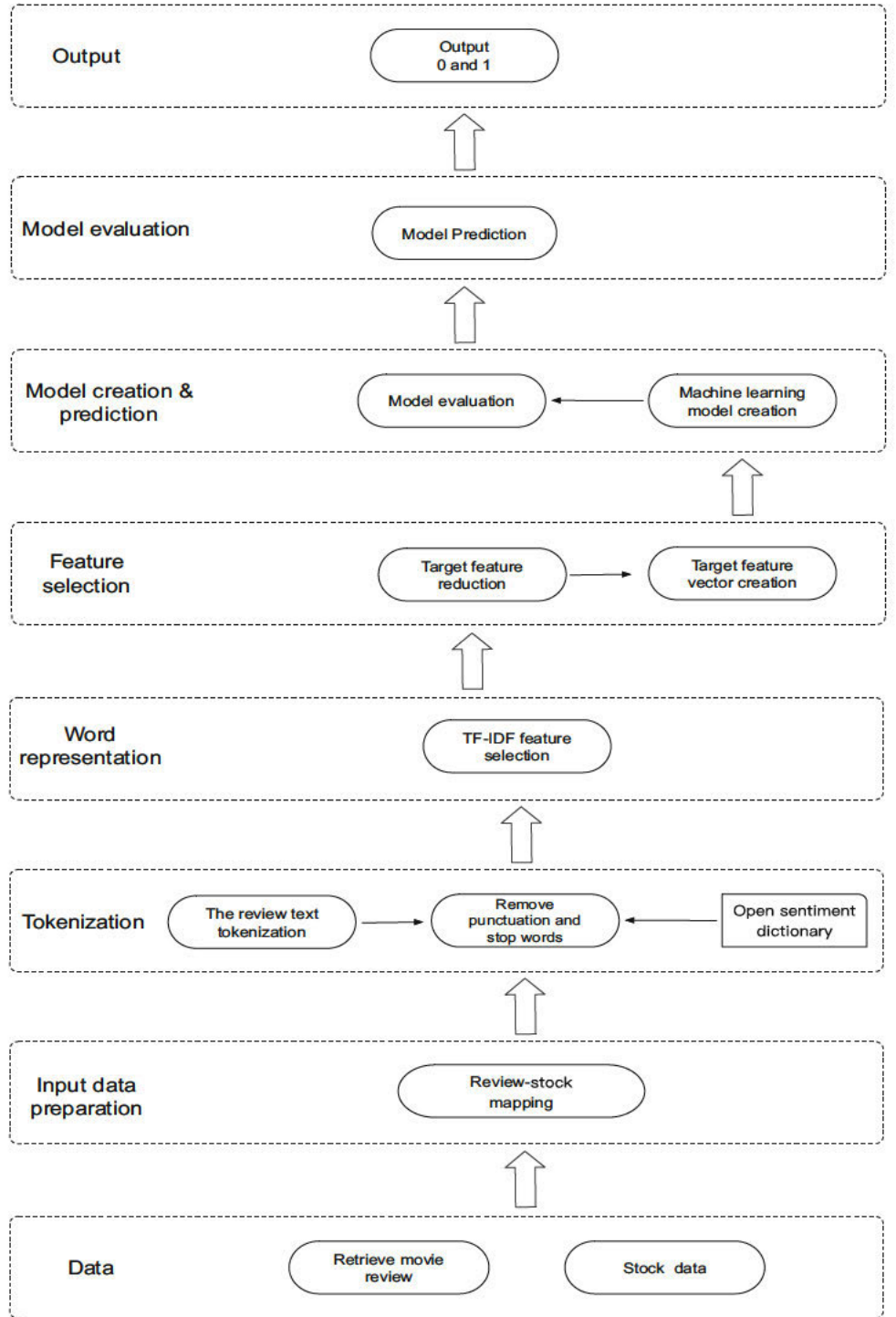
*Character attention* The words are composed of characters that represent different importance, and the contributions of the characters to indicate the overall sentiment polarity may vary. Se-attention is used in this attention network [61]. Hence, this architecture introduced HAN to distinguish different characters in a word.

$$u_{it} = \tanh(W_c h_{it}), \quad (16)$$

$$\alpha_{it} = \frac{\exp(u_{it})}{\sum_t \exp(u_{it})}, \quad (17)$$

$$V_i = \sum_t \alpha_{it} n_{it}, \quad (18)$$

**Fig. 3** The traditional machine learning workflow through TF-IDF feature extraction



where  $a_{it}$  considers audience reviews, and the importance of the  $t$ -th character in the word constructs the representation of the entire word  $V_i$ .  $W_c \in \mathbb{R}^v$  denotes the parameter for alignment function to calculate the match between  $h_{it}$  itself. We assess the importance of the character by the similarity  $u_{it}$  between the character level context vector  $u_c$ .

As briefly described above, the reason we leverage the character level attention is to find character-level emotional features and understand fine-grained eWOM features in audience sentiment.

## 2. Word level

**Word encoder** After obtaining the character vector  $S_i$ , this layer also adopts bidirectional LSTM to encode word



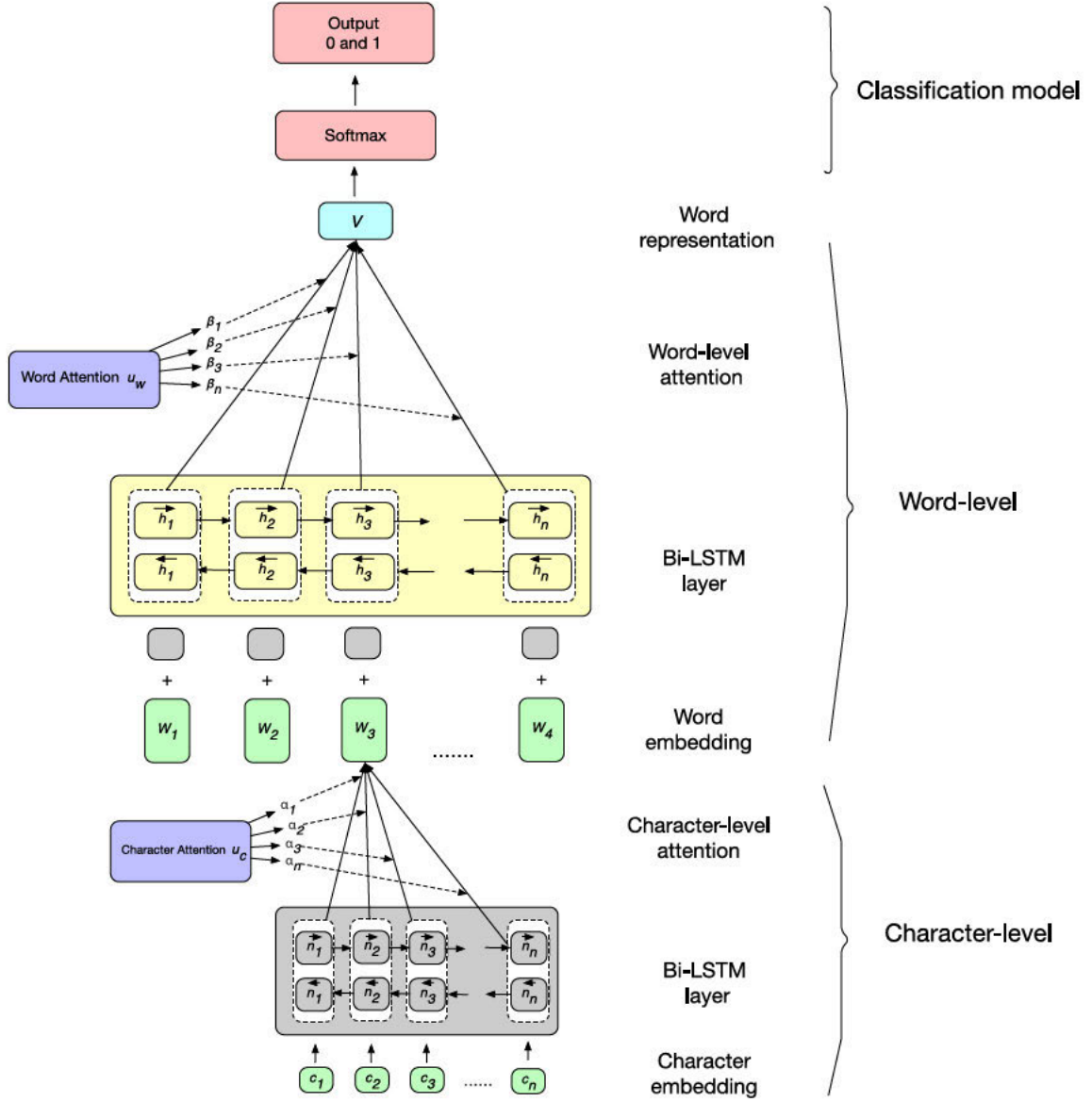


Fig. 4 A view of HTAN neural network

and gain a hidden layer representation  $h_i$  at the end of the time step. We gain the representation of word  $i$  through the concatenating of  $\vec{h}_i$  and  $\overleftarrow{h}_i$ ,

$$\vec{h}_i = \overrightarrow{\text{LSTM}}(x_i), \quad i \in [1, L], \quad (19)$$

$$\overleftarrow{h}_i = \overleftarrow{\text{LSTM}}(x_i), \quad i \in [L, 1], \quad (20)$$

$$h_i = [\vec{h}_i : \overleftarrow{h}_i]. \quad (21)$$

**Word attention** We then use a word-level context vector through attention mechanism to decode the word level representation among entire sentences. This attention layer assesses the importance of each word. This is denoted as follows:

$$u_{it} = \tanh(W_w h_i), \quad (22)$$

$$\beta_i = \frac{\exp(u_i)}{\sum_t \exp(u_i)}, \quad (23)$$

$$V = \sum_i \beta_i h_i, \quad (24)$$

The weight sum  $s$  is computed by  $V$ , and it can combine character-level attention with word-level attention in the review information. The effects of this attention layer are similar to those of the character level.

### 3. Classification model

In this model, we calculate the probability distribution for each category  $p$ .

$$p = \text{softmax}(W_s v + b_v), \quad (25)$$

where  $W_v$  and  $b_v$  present the weight parameters and offsets in the softmax layer. This outputs two classes (0 and 1); these denote positive or increase and negative or decrease, which reflect the availability of the attention model in studios. Finally, cross-entropy is used in a high-dimensional space as the training loss function.

$$L = - \sum_r \log p_{rj}, \quad (26)$$

where  $j$  is the label in audience review  $r$ .

#### 4.2.4 Hierarchical convoluted attention network (HCAN)

Based on the attention mechanism [16, 62], Fig. 5 shows our proposed model should pay more attention to the importance of review. Our framework should highlight the important sentiment characteristics of each review.

In our proposed HCAN architecture, first, we pre-process each review as input for the review corpus sequence on the word embedding layer, which encodes each review to a representation vector. In this way, these vectors fill into the convolutional layer and pooling layer, which generates the convolutional vector in each word. Second, the convolutional vectors are encoded by the bi-gated recurrent units (GRU). Finally, the attention mechanism layer assigns the attention value to each important word and calculates a weighted word vector, which represents the important feature for the entire context. The detail of the architecture is below.

**4.2.4.1 Embedding model** Word embedding places a word in a high-dimensional space map into a continuous vector space with a lower dimension. There are several types of word embedding trained, namely, word2vec [63] and Glove [64]. The Glove model, employed in our work, is a popular embedding algorithm in recent text mining for review analysis. This embedding layer converts each input word sequence into a corresponding embedding vector by consulting a lookup table.

Given the audience review  $r$  with  $n$  words containing  $w = \{w_1, w_2, w_3, \dots, w_n\}$ , we calculate each embedding vector through a word embedding layer. Each word in  $w$  is mapped into 100-dimensional vectors, which is looked up in pre-trained word embedding by Glove [64] for each word to denote the word embedding  $w_i = \{w_{1i}, w_{2i}, w_{3i}, \dots, w_{ni}\}$ . These word vectors input the convolutional layer.

**4.2.4.2 Convolutional model** The CNN is a combination of convolutional layer and pooling layer. The output of each layer is the input to the next layer. As a feature

extraction layer, the convolutional layer extracts word features through a filter, generating a feature map through a convolution kernel function. The pooling layer belongs to the feature mapping layer, which samples the feature word generated by the convolutional layer and outputs local optimal features. In this work, we introduced the feature representation of the convolution model, which is designed based on Text-CNN [65].

$$x_{1:n} = x_1 \oplus x_1 \oplus x_2 \oplus \dots \oplus x_n, \quad (27)$$

in which  $\oplus$  denote the concatenation [63]. The  $N$ -gram feature  $c_i$  represents the  $i$ -th eigenvalue in the word embedding, which is window size of  $X_{i:i+h-1}$ , which represents:

$$c_i = \sigma(wX_{i:i+h-1} + b), \quad (28)$$

where  $\sigma(\cdot)$  is convolution kernel function,  $w \in \mathbb{R}_{n \times k}$  is a filter,  $h = 5$  is the size of the sliding window, and  $b$  is the offset value.  $X_{i:i+h-1}$  means the feature matrix of word embedding from the  $i$ -th to the  $i + h - 1$ . Its length is 100. Hence, a feature map  $c$  denotes:

$$c = [c_1, c_2, \dots, c_{n-h+1}], \quad (29)$$

The characteristic table  $c$  is taken out as a feature of the next layer through a time maximization pooling operation.

$$c_{\max} = \begin{bmatrix} \max(c_1) \\ \vdots \\ \max(c_d) \end{bmatrix}, \quad (30)$$

Then, the max-pooled features are passed in next layer, and its length is also 100.

**4.2.4.3 Sequential mode** A document  $d$  with  $n$  words  $d_i = \{d_1, d_2, d_3, \dots, d_n\}$  merges word feature by CNN, and this new word vector is in the same position as  $w_i$ , which extracts crucial word features in the next attention layer. The review information is placed in the past or future by a bidirectional GRU, which calculates their potential feature vectors. We concatenate potential vectors from both directions for constructing a bidirectional encoded vector  $h_{is}$ , which follows as:

$$\overrightarrow{h_{is}} = \overrightarrow{\text{GRU}}(d_{is} : w_{is}), \quad i \in [1, L], \quad (31)$$

$$\overleftarrow{h_{is}} = \overleftarrow{\text{GRU}}(d_{is} : w_{is}), \quad i \in [L, 1], \quad (32)$$

$$h_{is} = [\overrightarrow{h_{is}}, \overleftarrow{h_{is}}], \quad (33)$$

In summary, we employ the attention mechanism for finding the most contributions of the sentence in the review corpus. Moreover,  $h_{is}$  inputs a single-layer perceptron for getting result  $u_{is}$ , which represents:

$$u_{is} = \tanh(W_w h_{is}), \quad (34)$$

To measure the importance of each review, we adopt  $\theta$  to express the similarity of a randomly initialized context

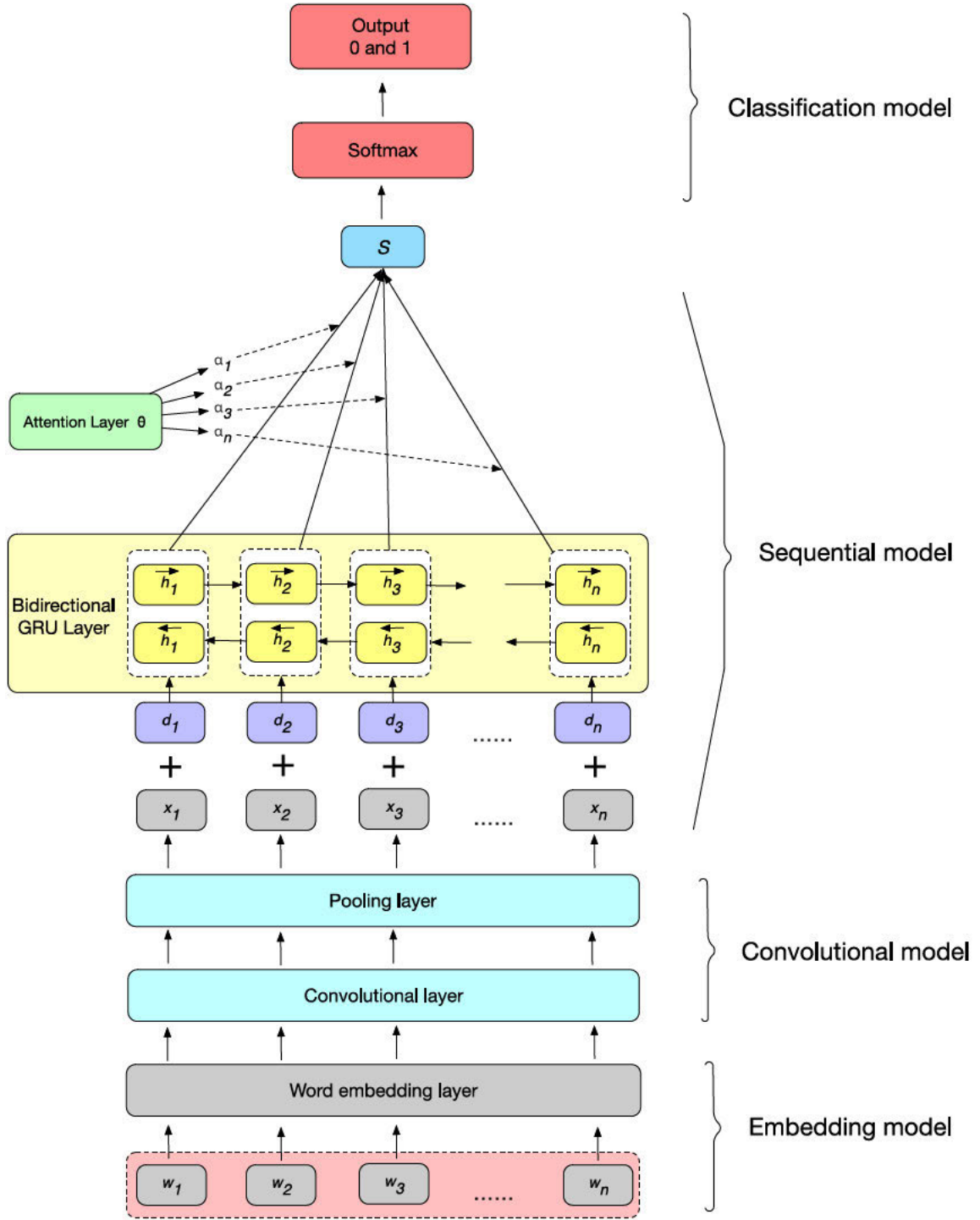


Fig. 5 A view of our HCAN neural network

vector  $u_i$ , and then we obtain a normalized attention weight matrix  $\alpha_i$  through softmax operation; this learning process denotes as:

$$\alpha_i = \frac{\exp(\theta_i u_{is})}{\sum_j \exp(\theta_j u_{js})}, \quad (35)$$

Furthermore, the weight value  $\alpha_i$  continues to compute the sum weight  $S$ ; its size is 100, and it shows that:

$$S = \sum_i \alpha_i h_{is}, \quad (36)$$

This approach combines sequential review by attention model, which facilitates to apply in the studio performance.



**4.2.4.4 Classification model** The softmax operation on two case study application: The output the probability distribution  $p$  over two categories,

$$p = (y = j|D) = \text{softmax}(W_s S + b_s), \quad (37)$$

where  $W_s$  is a weight matrix and  $b_s$  is bias vector that to be learned in the softmax layer. And  $j$  is the label of classification (0 and 1); these denote positive or increase and negative or decrease, respectively.

## 5 Experiment and results

### 5.1 Experiment setting

We compare the performance of traditional machine learning and deep learning. Table 6 presents the six algorithms used as a baseline in traditional machine learning, such as SVM, DT, LR, RF, AB and GB. Meanwhile, eight algorithms, CNN, LSTM, A-CNN, A-LSTM, UPA, HUAPA, HTAN, and HCAN, were relevant for deep learning in the two cases. The experiment results demonstrate that deep learning with the attention mechanism is superior to traditional machine learning by a bag-of-words approach. Furthermore, the result of the HCAN model fully illustrates the advantages of attention mechanisms and further improves the performance of deep learning. We constantly regulate the parameters and choose the optimal parameter values.

The baseline as traditional machine learning adopts Python with Scikit-learn in our experiment. Meanwhile, deep learning utilizes Python with TensorFlow 1.6 and Keras. The review corpus in our cases is large; moreover, the deep learning algorithm requires intensive computing. Our data pre-processing and model training was implemented with a computing system involving a 28-core Intel Xeon E5-2683 v3 CPU with 2.00 GHz and 64 GB RAM memory. This running system is Ubuntu Linux 16.04 64-bit version. Finally, we employ three metrics for the studio's performance accuracy, sensitivity, and specificity [66] in the two case studies.

### 5.2 Results

#### 5.2.1 Case study in eWOM classification

We report the result of the accuracy evaluation in the eWOM in Table 7. Additionally, we also compute the average of the seven studios for convenient comparison of each algorithm. Our model as HCAN achieved the best result, which was 88.70% average accuracy in the eight algorithms. We observe that the word embedding with the pre-trained Glove is superior to the feature extraction by

TF-IDF. The bag-of-words cannot capture semantic information in audience reviews, and when the word vocabulary is too large, it encounters a dimensional problem limitation. In fact, word embedding by Glove is a kind of distributed feature representation. Since no pre-trained word embedding is under the possibility of overfitting, we should choose word training by Glove. The capability of word embedding to represent textual information is stronger than that of the bag-of-words approach.

Compared to traditional machine learning, deep learning with our models achieved an average of almost 18% and 15% improvement, respectively. One of the main reasons may be the introduction of a caution mechanism that highlights the important parts of the review information. In addition, the CNN and LSTM algorithms contain implicit sequence information, which is strongly focused on parts that are similar to word embedding. It ignores the sentiment features in the long-distance words, which leads to less improvement in deep learning. We can also find that the average accuracy in A-CNN is 3.29% more than CNN. This result illustrates that deep learning incorporating attention mechanism is effective in capturing detailed features.

Our proposed model HCAN achieves the most advanced performance on this data set. It is worth noting that HCAN performed 2.17% better than HTAN, which indicates limited capability in character level attention. We believe this is because the importance of words depends not only on semantics but also on context and word order information. Extracting the features by the convolutional layer is more efficient than using only the character mechanism. Our model also considers the word features order and their relevance to sentiment words in the review. Our improvements prove that our model calculates attention weight of word from a global perspective by considering review information and incorporates important word vectors into the entire review information. Thus, these results illustrate excellent performance through deep learning with the attention mechanism. The attention mechanism model facilitates the classification of the sentiment polarity and further determines the emotional orientation of eWOM [67]. It also emphasizes the positive effect of emotional words in eWOM, which contributes to the enthusiasm of audience purchases.

#### 5.2.2 Case study in stock price movement

The results of the evaluation of stock price movement as shown in Table 8 are similar to the former case. An average of 67.43% for HCAN is the best result in these algorithms. HTAN obtains character-level embedding based on LSTM encoder in the context and computes embedding the weight between character level and word level. HCAN

**Table 6** The differences of models are used traditional machine learning and deep learning

Method	Algorithms
Traditional machine learning	Support vector machine (SVM) Decision tree (DT) Logistical regression (LR) Random forest (RF) AdaBoost (AB) Gradient Boosting (GB)
Deep learning	Convolutional neural network (CNN) Long short term memory (LSTM)
Attention based models	Attention based convolutional neural network (A CNN) Attention based long short term memory neural network (A LSTM) User and product attention neural network (UPA) Hierarchical user attention and product attention neural network (HUAPA)
Our models	Hierarchical two layer attention network (HTAN) Hierarchical convoluted attention network (HCAN)

**Table 7** Holistic comparison of traditional machine learning, deep learning and attention based model for eWOM, the performance is measured based on accuracy, these values are a percentage i.e., the higher the better

Algorithms	Studios							
	Disney	Fox	Lionsgate	Paramount	Columbia	Universal	Warner	Average
<i>Traditional machine learning</i>								
SVM	73.70	74.33	72.08	68.33	74.29	72.93	75.29	72.99
DT	68.15	69.26	64.78	64.32	69.55	66.05	68.78	67.27
LR	79.08	82.00	70.07	67.35	81.46	78.44	79.06	76.78
RF	72.34	76.34	69.16	65.96	74.15	72.66	73.50	72.02
AB	73.50	76.66	71.53	68.83	76.86	76.58	74.97	74.13
GB	71.79	77.22	70.07	65.38	78.48	75.83	74.73	73.36
<i>Deep learning</i>								
CNN	77.39	79.75	76.09	71.45	78.08	77.89	78.22	76.98
LSTM	77.88	80.76	73.18	72.03	81.33	77.27	77.93	77.20
<i>Attention based model</i>								
A CNN	83.06	81.04	79.85	79.49	79.20	84.62	82.17	80.27
A LSTM	80.82	76.99	73.87	79.63	79.63	80.55	80.55	77.09
UPA	85.37	83.10	84.16	83.99	82.44	84.56	83.81	83.38
HUAPA	83.81	82.21	80.48	83.97	83.97	85.41	86.88	83.06
<i>Our model</i>								
HTAN	85.57	84.29	89.12	84.23	86.71	89.42	86.34	<b>86.53</b>
HCAN	88.63	85.21	88.63	87.12	89.68	89.76	91.88	<b>88.70</b>

adopts a convolutional layer for the extraction feature strategy, which replaces the character level attention mechanism. Using bi-GRU encodes to guide the word level attention mechanism, each attention vector input corresponds to different semantics and emphasizes important words for stock price movement, such as *crazy*, *funny*,

*wonderful*, etc. Compared to average accuracy in both CNN and LSTM, the prediction accuracy in A-CNN and A-LSTM has 3.04% and 3.09% improvement, respectively. These results reflect the fine-grained emotional characteristics in the audience's review, which further promotes the positive growth of the stock market.

**Table 8** Holistic comparison of traditional machine learning, deep learning and attention based model for stock price movement, the performance is measured based on accuracy, these values are a percentage i.e., the higher the better

Algorithms	Studios							
	Disney	Fox	Lionsgate	Paramount	Columbia	Universal	Warner	Average
<i>Traditional machine learning</i>								
SVM	61.05	66.26	63.81	55.97	58.91	55.96	56.84	59.83
DT	61.72	63.90	61.98	55.65	59.91	63.16	53.26	59.94
LR	50.50	57.65	63.89	50.66	67.46	71.26	62.63	60.58
RF	66.15	64.04	64.42	61.07	65.05	62.07	53.54	62.33
AB	60.01	59.94	66.08	48.51	51.24	67.8	69.02	59.94
GB	58.17	71.05	71.10	50.65	68.52	55.57	67.04	63.16
<i>Deep learning</i>								
CNN	58.67	66.89	53.7	53.59	53.79	65.54	72.8	60.72
LSTM	53.15	67.12	65.52	59.16	63.2	59.5	60.36	61.16
<i>Attention based model</i>								
A CNN	62.65	62.65	57.23	57.17	55.84	69.44	74.91	63.76
A LSTM	56.34	70.68	68.99	61.32	67.12	61.69	63.64	64.25
UPA	63.64	68.3	68.30	59.25	61.22	69.62	68.92	63.67
HUAPA	63.75	70.36	70.36	64.21	72.90	60.86	64.05	65.84
<i>Our model</i>								
HTAN	58.24	67.51	58.69	61.32	66.59	69.81	62.94	<b>63.59</b>
HCAN	71.15	70.03	60.58	67.10	78.68	60.02	64.45	<b>67.43</b>

The impact of the attention mechanism in the stock price movement is that the characters can be captured by the word-level attention layer. The word-level attention mechanism mainly detects different fine-grain features in audience reviews and defines the value of each review on the sentiment of eWOM. The word-level attention mechanism with the feature extracted by convolutional layers improves model performance compared to models that use only character-level or word-level attention mechanisms. These results demonstrate that review features are reflected in different semantic levels, which presents the impact of audience review on eWOM at different levels, and further illustrates the impact of eWOM on studio performance.

### 5.2.3 Additional analysis for the two case studies

**5.2.3.1 Model analysis: effect of attention mechanisms** In Tables 7 and 8, comparison between deep learning model and attention model in two case studies, UPA and HUAPA both achieve some improvements, which proves the rationality of incorporating eWOM and stock market via attention mechanisms. These results also illustrate that attention mechanisms can capture more eWOM feature in the audience reviews.

Compared to HTAN and HCAN, word-level features are more effective than character-level features to enhance

audience reviews representation. Although some character in words can show the eWOM feature, which is finally determined by the audience sentiment. Hence, it is reasonable that eWOM feature of the audience is most apparent at the word level.

Compared to the attention-based model (e.g., A-CNN, A-LSTM, UPA and HUAPA), our models achieved better performance, which indicates that both the character level and word level information can contribute to studios performance. The results also demonstrate that HTAN and HCAN can capture the audience interest, which provides decision support for the studios.

**5.2.3.2 Robustness test** To assess the effectiveness of a robustness check of each algorithm in Tables 9 and 10, we adopt the sensitivity and specificity for the metric of additional testing. The results confirm our findings that these models are robust for small noise in the review corpus and are not particularly impacted by slight multicollinearity. The average sensitivity for HCAN achieves 92.82% and 76.05%, respectively. Overall, using deep learning with attention mechanism helps to improve studio performance.

**5.2.3.3 Ablation study** The ablation study results shown in Table 11 indicate that when we reduce the attention



**Table 9** Comparison of sensitivity (Sens.) and specificity (Spec.) in eWOM, with values presented as percentage, i.e., the higher the better

Method	Disney Sens./Spec.	Fox Sens./Spec.	Lionsgate Sens./Spec.	Paramount Sens./Spec.	Columbia Sens./Spec.	Universal Sens./Spec.	Warner Sens./Spec.
<i>Traditional machine learning</i>							
SVM	62.21/79.01	57.04/80.94	43.62/86.94	50.77/81.49	51.10/81.87	56.29/79.73	53.89/85/64
DT	52.44/77.99	42.96/79.33	29.79/93.06	50.96/74.32	42.86/78.28	45.37/74.49	49.02/78.34
LR	56.47/93.24	46.01/95.78	14.90/98.89	28.35/96.56	32.97/97.31	35.15/96.12	44.64/95.70
RF	50.47/86.04	38.46/90.83	20.74/94.44	36.60/87.95	29.12/88.87	32.54/89.04	36.74/91.29
AB	49.52/88.51	35.12/92.56	30.32/93.06	40.61/89.96	36.81/89.95	38.48/92.14	42.14/90.82
GB	32.90/96.14	25.54/97.00	14.89/98.89	23.94/96.41	24.73/96.05	26.37/96.02	29.20/96.76
<i>Deep learning</i>							
CNN	80.89/71.78	91.17/49.93	82.78/63.30	88.38/48.85	98.92/14.29	86.61/56.53	88.76/56.45
LSTM	89.53/59.26	90.72/54.72	90.00/40.96	85.22/54.41	96.41/35.16	86.61/54.39	87.64/57.54
<i>Attention based model</i>							
A CNN	93.02/58.81	91.3/55.71	91.96/41.5	94.24/44.05	92.2/44.51	96.47/44.77	95.3/44.64
A LSTM	86.44/65.54	84.81/57.68	94.95/26.61	91.2/36.59	95.11/21.35	93.47/39.47	93.47/39.47
UPA	88.52/36.97	88.32/39.77	85.31/40.58	92.69/33.17	88.62/35.39	94.37/33.98	97.06/40.26
HUAPA	87.12/39.38	87.38/44.66	94.78/25.13	92.41/24.27	92.41/24.27	93.72/25.76	91.94/37.19
<i>Our model</i>							
HTAN	85.01/16.13	86.24/35.36	78.86/38.55	92.21/22.36	85.58/26.47	92.98/24.56	98.74/35.66
HCAN	87.65/14.69	88.57/30.44	93.58/22.65	94.57/10.98	95.88/33.78	94.85/14.21	94.63/21.78

**Table 10** Comparison of sensitivity (Sens.) and specificity (Spec.) in stock price movement, with values presented as a percentage, i.e., the higher the better

Method	Disney Sens./Spec.	Fox Sens./Spec.	Lionsgate Sens./Spec.	Paramount Sens./Spec.	Columbia Sens./Spec.	Universal Sens./Spec.	Warner Sens./Spec.
<i>Baseline: traditional machine learning</i>							
SVM	59.75/53.77	59.25/57.37	54.7/58.01	56.67/54.18	54.53/54.69	61.07/56.31	55.45/54.59
DT	53.35/52.81	57.80/59.18	50.12/56.96	53.38/55.35	54.53/54.69	54.38/57.64	54.03/50.52
LR	54.0/55.22	59.42/62.69	61.82/56.85	53.48/55.32	53.75/57.96	59.06/57.53	61.89/53.81
RF	62.18/52.53	66.08/59.86	60.83/58.51	52.32/51.40	60.85/52.23	59.06/57.53	53.29/54.25
AB	60.48/55.02	59.14/59.95	61.80/58.58	56.32/48.66	54.01/59.09	59.29/54.38	58.77/54.13
GB	60.48/55.02	58.28/60.56	62.93/58.12	52.25/54.16	55.01/55.07	54.41/53.64	56.7/55.20
<i>Deep learning</i>							
CNN	74.45/26.75	27.16/89.77	90.41/16.75	60.88/46.17	62.85/51.56	68.27/45.48	50.58/45.05
LSTM	58.63/53.72	95.32/20.23	67.15/34.92	56.6/52.52	72.35/41.33	67.88/47.71	77.27/29.36
<i>Attention based model</i>							
A CNN	76.81/30.44	28.76/92.77	28.76/92.77	62.71/49.58	65.83/54.17	70.4/50.09	52.95/49.2
A LSTM	59.87/56.89	97.21/23.98	68.4/40.02	58.91/55.99	73.58/46.04	69.94/52.63	78.63/35.23
UPA	78.63/35.23	49.41/74.60	90.73/29.5	67.56/49.33	73.66/51.10	74.43/49.62	74.43/49.62
HUAPA	69.56/55.06	81.6/36.71	73.9/46.28	67.2/53.19	74.76/51.77	74.53/54.18	77.91/40.60
<i>Our model</i>							
HTAN	74.02/50.15	70.06/56.43	88.74/38.1	72.41/49.08	81.5/48.03	78.45/49.16	74.99/53.99
HCAN	79.26/53.23	65.99/49.44	79.39/52.54	75.50/50.39	75.94/57.51	79.12/55.73	77.18/45.98

layer in HCAN at two levels, we observed 80.45% in “- Word level attention” and 83.36% in “- Character level convolution,” lower than results from the original method.

The results fully demonstrate the great contribution of the attention mechanism in these two cases. This result also proves the contribution of character level [55, 56]. In

**Table 11** Ablation test of HTAN and HCAN, where the values are presented as percentages

Cases	HCAN		HTAN	
eWOM	Original	88.70	Original	86.53
	Character level convolution	83.36	Character level attention	84.27
	Word level attention	80.45	Word level attention	77.67
			+ Word level convolution	<b>88.57</b>
Stock price movement	Original	67.43	Original	63.59
	Character level convolution	63.80	Character level attention	60.12
	Word level attention	64.21	Word level attention	60.95
			+ Word level convolution	<b>65.80</b>

“Character level attention” means that an attention layer is added at the character level. Similarly, “+ word level convolution” means a convolutional layer was added at the word level, and “- word level convolution” refers to a subtract convolutional layer at the word level

addition, we found that the performance of HCAN is better than the performance of HAN when covering the convolution layer. Obviously, the accuracy 88.57% and 68.8% in “+ Word level convolution” in two cases are better than other layers. Even though HAN operates sequential encoding at the character level and convolution layer at the word level, this operation process is HCAN, which outperforms both the HAN and hierarchical convolution network (HCN). These results also illustrate that the convolution layer can capture locality information of audience reviews, which proves important emotional information is related to a crucial sentiment word.

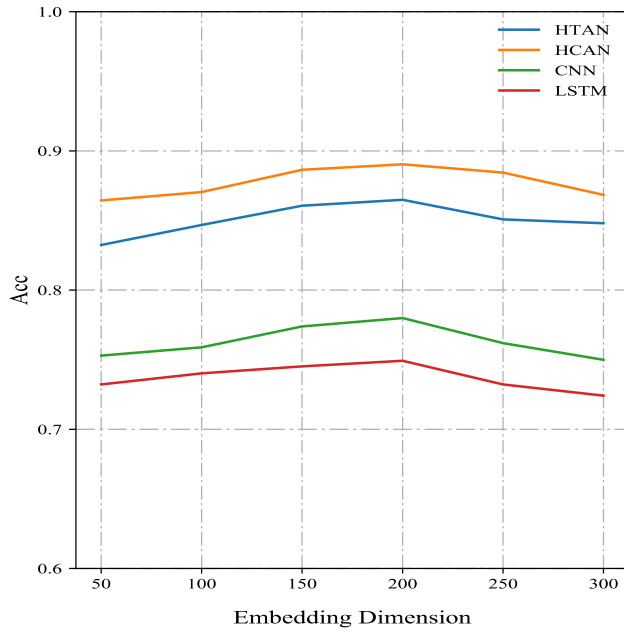
Previous work [68] had proved that dimensionality of embedding can influence the performance of neural model. To assess the effect of potential word dimensions, we compare the eWOM performance in different dimensions between embedding layer and attention layer, respectively. In Fig. 6a, the size of the word embedding dimension set to 50, 100, 150, 200, 250, and 300, respectively. As shown, the accuracy of HCAN is significantly higher than that of other algorithms. Furthermore, in Fig. 6b, the word dimension interval is reduced under the attention dimension. We also observed that the accuracy of HCAN outperforms that of HTAN. The main advantage of HCAN over HTAN is that CNN employed at character level can capture accurate contextual information in audience reviews through convolutional kernels and word embedding.

**5.2.3.4 Qualitative analysis in attention mechanism** The audience review for the prediction of the eWOM and the prediction of the stock market price through eWOM is a fine-grained task [21]; it is necessary to analyze which word decides the sentiment polarity. We conducted a case study by visualizing the related attention weights, based on several sample sentences picked from the test set. The attention weight  $u_i$  in our proposed model HCAN can

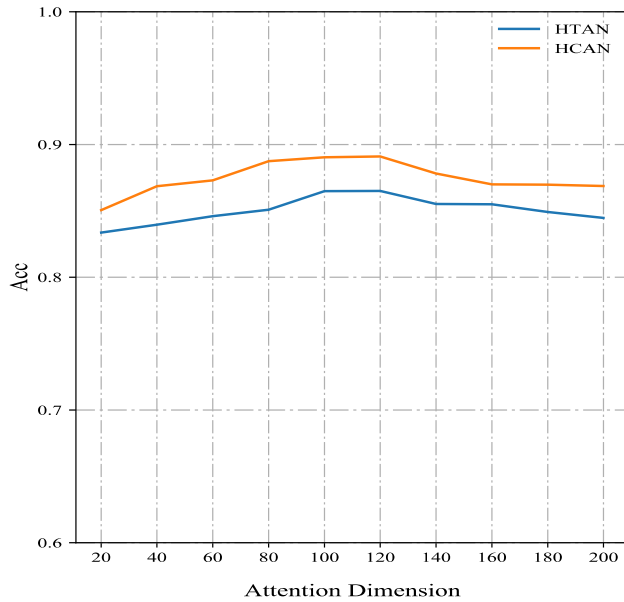
visualize the attention weights accordingly. Figure 7 shows that representation of how attention is focused on words with critical influence. We adopt the Heml [69] tool to visualize each sentence of the review corpus. The dark color expresses that the weight in attention mechanism is more important, and the ordinate represents the label of the classification.

The sentence in Fig. 7a “*I went in with no expectation and was delightfully surprised*” is the corresponding positive eWOM. Evidently, the attention mechanism can dynamically detect the important parts of sentence semantics. We also clearly observed that the word “*surprised*” is the darkest color, which demonstrates that “*surprised*” is the biggest contributor to the performance of positive eWOM. In Fig. 7b “*Worst Star Wars movie, I have ever seen,*” the word color of “*worst*” is the darkest. This sentence clearly expresses the audience is disappointed in *Star Wars* by the negative eWOM. Thus, the positive eWOM proves that the movie is more exciting than the audience expected, which increases movie sales. “*Watching this movie was fun! Isn’t that what going to the movies is all about with this genre of film?*” in Fig. 7c illustrates that this audience is very interested in this movie, which promotes to increase the stock price of the studios. Conversely, the word “*disaster*” is the depth color in Fig. 7d, which indicates the audience expressed meaning is the opposite of Fig. 7c. This also provides the evidence that positive eWOM causes the stock price of studios to increase, and negative eWOM causes the stock price of the studios to decrease. All in all, these visualized results help to explain why our model significantly outperforms the baseline systems.

We further analyze the capability of our attention model for discovering informative words. The hot word frequency is counted in the top three word-weights through the attention mechanism in Fig. 8. We observe a lot of positive words in the task of the eWOM in Fig. 8a, for instance, 385



(a)



(b)

Fig. 6 Validation accuracy is in eWOM

times for “good” and 380 times for “great.” The opposite result, 242 times for “bad,” appears most often in the prediction task of stock price movement in Fig. 8b, which expounds that the eWOM is negative with the motion picture industry and leads to the decline of the stock price of the studios. Exciting is that the word “best” appears the second highest amount of times in Fig. 8b, which indicates that most of the audience is satisfied with the market

performance of the current movies. This prompts the studios to produce better movies and increase their stock income. Overall, the online review task analysis through attention mechanism is very important and effective. Because there are massive reviews on IMDb, it is important to accurately grasp the eWOM direction, which improves the capability of the movie marketing and expands the studio’s performance.

**5.2.3.5 Errors analysis** In the first category of Table 12, this example has two emotional terms, “really fun” and “ridiculous.” The proposed method assigns large attention weights to words like “ridiculous.” We find that the attention weight of “really fun” is less than the attention weight of “ridiculous”; thus, the eWOM label of this sentence is predicted incorrectly. Moreover, we also observe that the word “crazy” is utilized in two different sentences, which explains the different emotional polarity. The first sentence means the audience is excited or even crazy. In the second sentence, the word “stupid” also occurs, but our model catches the error polarity. Therefore, the general model has difficulty captured the implicit eWOM in complex semantics.

In the fourth category in stock price movement, the sentences are objective, which is without any expression of opinion. It is difficult to deal with such types of errors using the deep learning technique; especially the low frequency of these expressions makes it difficult for the attention mechanism to capture their patterns. Hence, how to better understand the eWOM polarity of audience reviews is still a challenge in deep learning with attention mechanism.

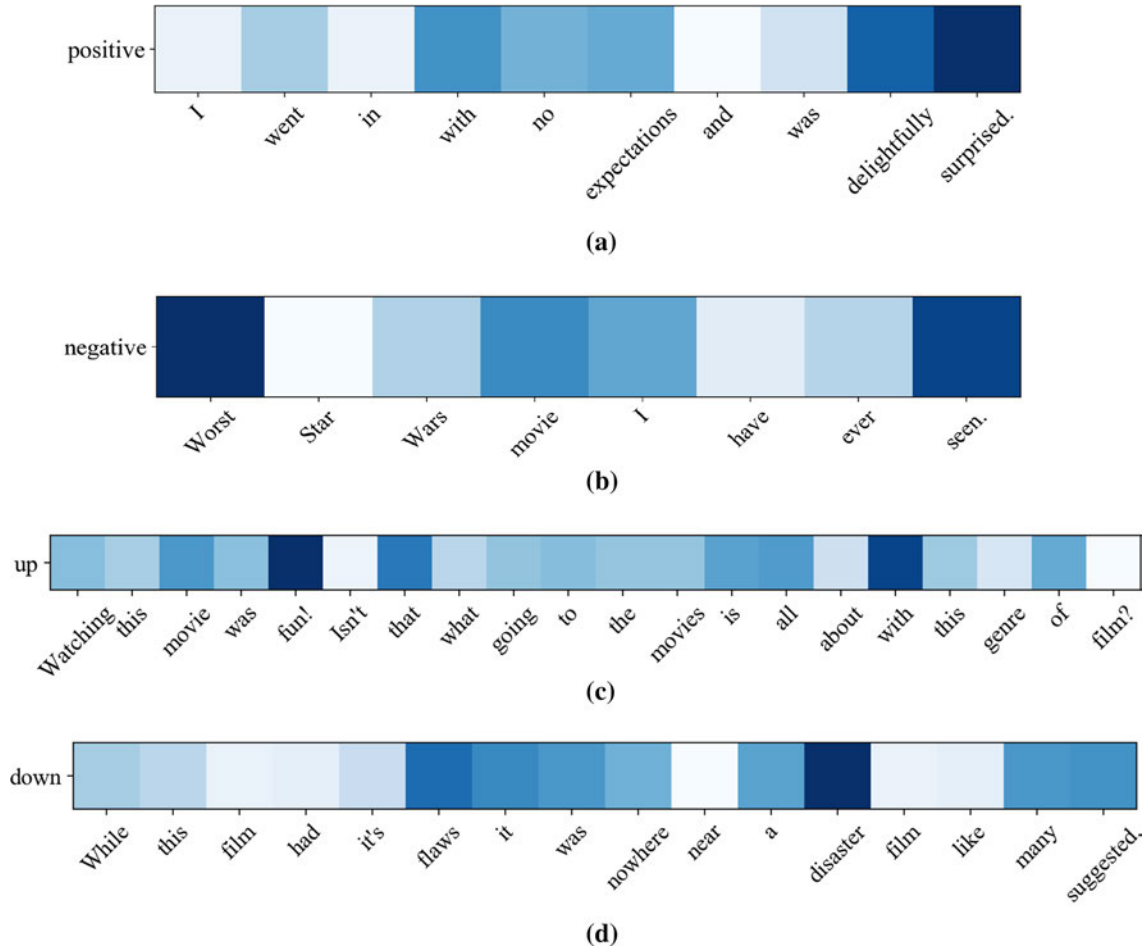
## 6 Discussion

### 6.1 Implications for business research and practice

This study should establish a new assessment plan of eWOM for the movie studio through deep learning with attention mechanism. More specifically, in the US movie market as a research scenario, we further investigate eWOM and stock market movement in Hollywood studios. More importantly, we assert that the performance of the studios isn’t an isolated occurrence, but rather exists as a collective presence using eWOM. Therefore, we believe that the decision making of the studio’s benefits from the decision support by the eWOM feature.

Even though attention mechanisms have recently gained considerable appeal in academia, their application in the industry is still rare. The possible reason is their complexity and uncertainty. In addition, the eWOM of audience reviews also presents difficulties for the attention





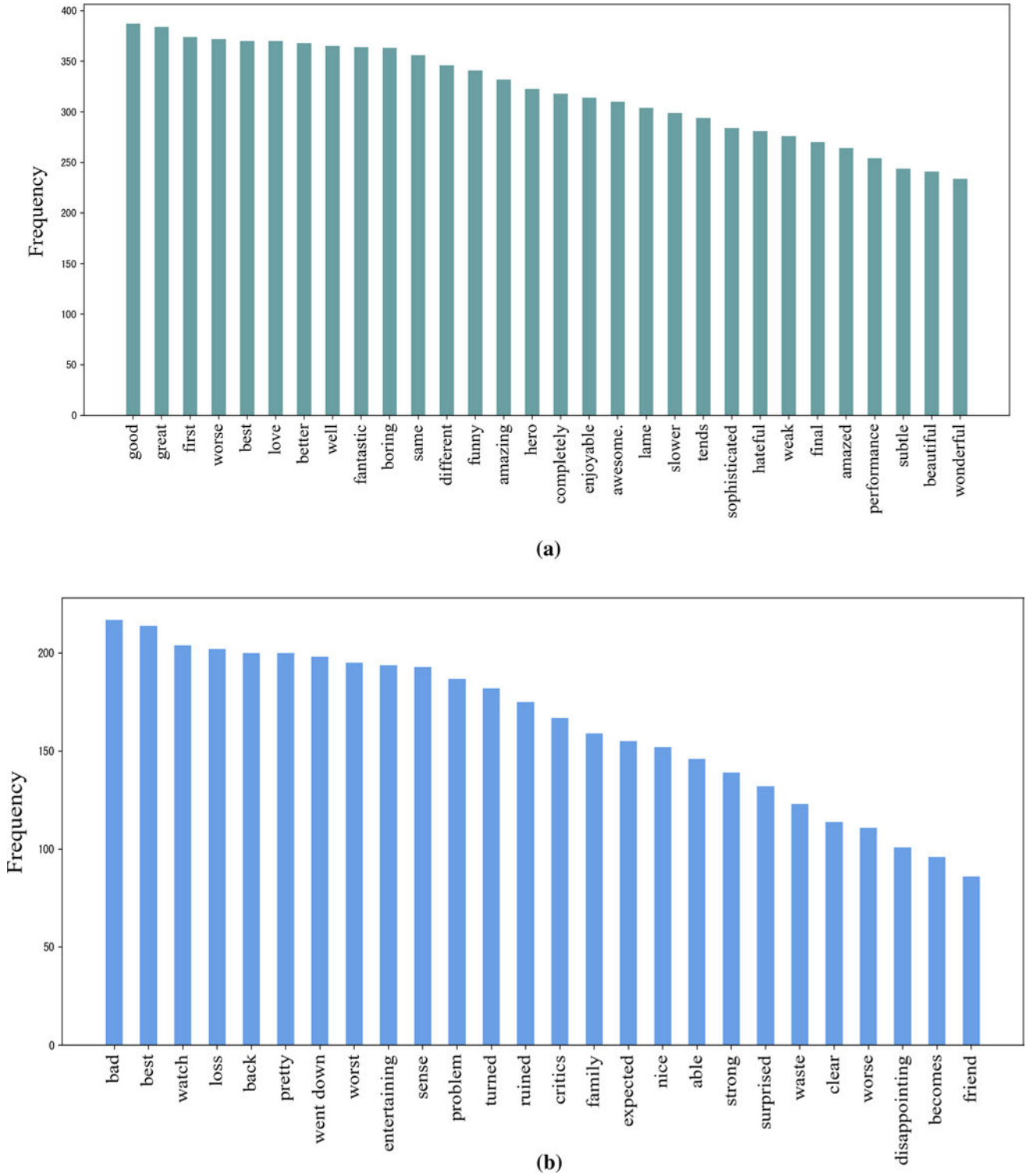
**Fig. 7** Attention visualizations. **a, b** Classification of the eWOM, **c, d** prediction the stock price movement

mechanism model. Since audience reviews have become the main source of movie rating, studios must understand the emotional needs of the audience. These results were consistent with previous research [7, 8]. This study extends early research through the application of attention mechanisms, which examines the effect of the eWOM features of audience reviews. This underscores the need for movie practitioners to have a comprehensive understanding of deep learning in order to benefit from the attention mechanism.

The effect of computing eWOM through attention mechanism provides new opportunities for the business model of the studios. Here, eWOM of audience reviews by attention mechanism is still an important form of communication. It also attempts to apply to speech recognition, intelligent customer service, image recognition, and machine translation [70]. Finally, our attention-based framework can support the existing industry chain; we also believe that traditional models will soon be replaced by deep architecture.

## 6.2 Limitations and potential for future research

Our research has some limitations that can facilitate future research. First, like other decision supports that use audience review, the black box of deep learning with attention mechanisms is difficult to interpret. Opening the trained black box can apply in two cases. Second, our research uses only IMDb as the source of review data by eWOM. Future research can investigate other sources, such as news disclosure and market reports. Third, many of the latest deep learning models have not been explored in our research. Such as transformer attention [61] and bidirectional encoder representations from transformers (BERT) [71]. Fourth, we did not consider the impact of special movies on the stock market of the studios, for instance, the simultaneous release of movies and the impact of pre-screen advertising. Fifth, we only considered the same day to match between audience review and the stock price movement. We thereby will closely match the time



**Fig. 8** The statistics of the hot word frequency. **a** The hot word is applied in the eWOM, **b** the hot word is applied in the prediction of the stock market price

granularity in the future. Finally, we can incorporate other aspects, such as audience satisfaction.

In future work, we will overcome the above limitations and propose more effective application frameworks for

decision support for studios. We believe that our research is a potentially useful first step using deep learning with attention mechanisms, which verifies the practicality of the attention mechanism model through audience reviews.

**Table 12** Example for each error category

	Errors category	Examples	Ture label	Predictive label
eWOM	Incorrect attention distribution	I had a really fun time in the cinema watching this. but I feel people are hyping this to ridiculous levels	1	0
	Double meaning words	I see most reviews separate the audience into two; those who have seen the original anime series and those who haven't. That was crazy ok!	1	0
		First time reviewer here but I have to do this because I feel nearly majority of the reviews here are stupid and crazy	0	1
Stock price movement	Hard to recognize	Deadpool is directed by Tim Miller and stars Ryan Reynolds, Morena Baccarin, T.J. Miller and Ed Skrein	1	0

## 7 Conclusions

Since massive data on audience review are injected into the movie market every day, managers, marketers and researchers need smarter models to process this information. We introduced deep learning in two cases of studios, which consists of eWOM and stock market of studios. We establish a review corpus from a Hollywood studio, and our model HCAN achieves superior predictive performance by inferring important information from an ordered sequence of comments and captures highly nonlinear relationships. As a consequence, eWOM by incorporating attention mechanisms into deep learning models can provide an excellent insight into the movie market.

**Acknowledgements** This work is supported by the China Scholarship Council (CSC) throughout the Grant Number: 201508390019.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Piryan R, Madhavi D, Singh VK (2017) Analytical mapping of opinion mining and sentiment analysis research during 2000–2015. *Inf Process Manag* 53:122–150. <https://doi.org/10.1016/j.ipm.2016.07.001>
- Mäntylä MV, Graziotin D, Kuuttila M (2018) The evolution of sentiment analysis – a review of research topics, venues, and top cited papers. *Comput Sci Rev* 27:16–32. <https://doi.org/10.1016/j.cosrev.2017.10.002>
- Schlosser AE (2011) Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. *J Consum Psychol* 21:226–239. <https://doi.org/10.1016/j.jcps.2011.04.002>
- Chu S C, Kim Y (2011) Determinants of consumer engagement in electronic word of mouth (eWOM) in social networking sites. *Inte J Advert* 30:47–75. <https://doi.org/10.2501/IJA.30.1.047.075>
- Gretzel U, Yoo KH (2008) Use and impact of online travel reviews. In: O'Connor P, Höpken W, Gretzel U (eds) *Information and communication technologies in tourism 2008*. Springer, Vienna, pp 35–46
- Park C, Lee TM (2009) Information direction, website reputation and eWOM effect: a moderating role of product type. *J Bus Res* 62:61–67. <https://doi.org/10.1016/j.jbusres.2007.11.017>
- Duan W, Gu B, Whinston A (2008) The dynamics of online word of mouth and product sales – an empirical investigation of the movie industry. *J Retail* 84:233–242. <https://doi.org/10.1016/j.jretai.2008.04.005>
- Liu Y (2006) Word of mouth for movies: its dynamics and impact on box office revenue. *J Mark* 70:74–89. <https://doi.org/10.1509/jmkg.70.3.74>
- Kao YC, Shyu J, Huang JY (2015) eWOM for stock market by big data methods. *J Acc Finance Manage Strategy* 10(2):93–106
- Luo X, Zhang J (2013) How do consumer buzz and traffic in social media marketing predict the value of the firm? *J Manag Inf Syst* 30:213–238. <https://doi.org/10.2753/MIS0742.1222300208>
- Mishra A, Satish SM (2016) eWOM: extant research review and future research avenues. *Vikalpa* 41:222–233. <https://doi.org/10.1177/0256090916650952>
- Blei DM (2003) Latent dirichlet allocation. *J Mach Learn Res* 3:993–1022
- Lee DD, Seung HS (2000) Algorithms for non negative matrix factorization. *Adv Neural Inf Process Syst* 2001:556–562
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
- Greene S, Resnik P (2009) More than words: syntactic packaging and implicit sentiment. *Association for Computational Linguistics*, Boulder, Colorado, p 503
- Bahdanau D, Cho K, Bengio Y (2014) Neural machine translation by jointly learning to align and translate. [arXiv:1409.0473\[cs,stat\]](https://arxiv.org/abs/1409.0473)
- Chaudhari S, Polatkan G, Ramanath R, Mithal V (2019) An attentive survey of attention models. [arXiv:1904.02874\[cs,stat\]](https://arxiv.org/abs/1904.02874)
- Chen H, Sun M, Tu C et al (2016) Neural sentiment classification with user and product attention. *Association for Computational Linguistics*, Austin, pp 1650–1659
- Yang Z, Yang D, Dyer C et al (2016) Hierarchical attention networks for document classification. *Association for Computational Linguistics*, pp 1480–1489
- Yin W, Schütze H, Xiang B, Zhou B (2016) ABCNN: attention based convolutional neural network for modeling sentence pairs. *Trans Assoc Comput Linguist* 4:259–272. <https://doi.org/10.1162/tacl.2016.00097>
- Wang Y, Huang M, Xiaoyan Z, Zhao L (2016) Attention based LSTM for aspect level sentiment classification. *Association for Computational Linguistics*, Austin, pp 606–615

22. Wu Z, Dai X Y, Yin C et al (2018) Improving review representations with user attention and product attention for sentiment classification. [arXiv:180107861\[cs\]](https://arxiv.org/abs/180107861)
23. Dellarocas CN, Awad N, Zhang X (2004) Using online reviews as a proxy of word of mouth for motion picture revenue forecasting. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.620821>
24. Chakravarty A, Liu Y, Mazumdar T (2010) The differential effects of online word of mouth and critics' reviews on pre release movie evaluation. *J Interact Mark* 24:185–197. <https://doi.org/10.1016/j.intmar.2010.04.001>
25. Kim SH, Park N, Park SH (2013) Exploring the effects of online word of mouth and expert reviews on theatrical movies' box office success. *J Media Econ* 26:98–114. <https://doi.org/10.1080/08997764.2013.785551>
26. Ullah R, Zeb A, Kim W (2015) The impact of emotions on the helpfulness of movie reviews. *J Appl Res Technol* 13:359–363. <https://doi.org/10.1016/j.jart.2015.02.001>
27. Hur M, Kang P, Cho S (2016) Box office forecasting based on sentiments of movie reviews and Independent subspace method. *Inf Sci* 372:608–624. <https://doi.org/10.1016/j.ins.2016.08.027>
28. Lash MT, Zhao K (2016) Early predictions of movie success: the who, what, and when of profitability. *J Manag Inf Syst* 33:874–903. <https://doi.org/10.1080/07421222.2016.1243969>
29. Lee JH, Jung SH, Park J (2017) The role of entropy of review text sentiments on online WOM and movie box office sales. *Electron Commer Res Appl* 22:42–52. <https://doi.org/10.1016/j.elerap.2017.03.001>
30. Baek H, Oh S, Yang H D, Ahn J (2017) Electronic word of mouth, box office revenue and social media. *Electron Commer Res Appl* 22:13–23. <https://doi.org/10.1016/j.elerap.2017.02.001>
31. Xun J, Guo B (2017) Twitter as customer's eWOM: an empirical study on their impact on firm financial performance. *Int Res* 27:1014–1038. <https://doi.org/10.1108/IntR-07-2016-0223>
32. Vermeer SAM, Araujo T, Bernitter SF, van Noort G (2019) Seeing the wood for the trees: how machine learning can help firms in identifying relevant electronic word of mouth in social media. *Int J Res Mark*. <https://doi.org/10.1016/j.ijresmar.2019.01.010>
33. Kim T, Kim WG, Kim HB (2009) The effects of perceived justice on recovery satisfaction, trust, word of mouth, and revisit intention in upscale hotels. *Tour Manag* 30:51–62. <https://doi.org/10.1016/j.tourman.2008.04.003>
34. Yin D, Bond SD, Georgia Institute of Technology et al (2014) Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Q* 38:539–560. <https://doi.org/10.25300/MISQ/2014/38.2.10>
35. Luo X (2007) Consumer negative voice and firm idiosyncratic stock returns. *J Mark* 71(3):75–88
36. Fama EE (1970) Efficient capital markets: a review of theory and empirical work. *J Finance* 25:36
37. Li Q, Chen Y, Wang J et al (2018) Web media and stock markets : a survey and future directions from a big data perspective. *IEEE Trans Knowl Data Eng* 30:381–399. <https://doi.org/10.1109/TKDE.2017.2763144>
38. Tetlock PC (2007) Giving content to investor sentiment: the role of media in the stock market. *J Finance* 62(3):1139–1168
39. Li Q, Wang T, Li P et al (2014) The effect of news and public mood on stock movements. *Inf Sci* 278:826–840. <https://doi.org/10.1016/j.ins.2014.03.096>
40. Nguyen TH, Shirai K, Velcin J (2015) Sentiment analysis on social media for stock movement prediction. *Expert Syst Appl* 42:9603–9611. <https://doi.org/10.1016/j.eswa.2015.07.052>
41. Li Q, Wang T, Gong Q et al (2014) Media aware quantitative trading based on public Web information. *Decis Support Syst* 61:93–105. <https://doi.org/10.1016/j.dss.2014.01.013>
42. Si J, Mukherjee A, Liu B et al (2014) Exploiting social relations and sentiment for stock prediction. In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. Association for Computational Linguistics, Doha, Qatar, pp 1139–1145
43. Cheng J, Zhao S, Zhang J et al (2017) Aspect level sentiment classification with HEAT (HiErarchical ATtention) network. In: *Proceedings of the 2017 ACM on conference on information and knowledge management CIKM'17*. ACM Press, Singapore, pp 97–106
44. Fan W, Gordon MD (2014) The power of social media analytics. *Commun ACM* 57:74–81. <https://doi.org/10.1145/2602574>
45. Xing FZ, Cambria E, Welsch RE (2018) Natural language based financial forecasting: a survey. *Artif Intell Rev* 50:49–73. <https://doi.org/10.1007/s10462-017-9588-9>
46. Kim T, Hong J, Kang P (2015) Box office forecasting using machine learning algorithms based on SNS data. *Int J Forecast* 31:364–390. <https://doi.org/10.1016/j.ijforecast.2014.05.006>
47. Du J, Xu H, Huang X (2014) Box office prediction based on microblog. *Expert Syst Appl* 41:1680–1689. <https://doi.org/10.1016/j.eswa.2013.08.065>
48. Henrique BM, Sobreiro VA, Kimura H (2019) Literature review: machine learning techniques applied to financial market prediction. *Expert Syst Appl* 124:226–251. <https://doi.org/10.1016/j.eswa.2019.01.012>
49. Ding X, Zhang Y, Liu T, Duan J (2015) Deep learning for event driven stock prediction. In: *Twenty fourth international joint conference on artificial intelligence*
50. Kraus M, Feuerriegel S (2017) Decision support from financial disclosures with deep neural networks and transfer learning. *Decis Support Syst* 104:38–48. <https://doi.org/10.1016/j.dss.2017.10.001>
51. He R, Lee WS, Ng HT, Dahlmeier D (2018) Exploiting document knowledge for aspect level sentiment classification. [arXiv:180604346\[cs\]](https://arxiv.org/abs/180604346)
52. Liu J, Zhang Y (2017) Attention modeling for targeted sentiment. *Association for Computational Linguistics, Valencia*, pp 572–577
53. Hu Z, Liu W, Bian J et al (2018) Listening to chaotic whispers: a deep learning framework for news oriented stock trend prediction. *ACM Press, New york*, pp 261–269
54. Liu Q, Cheng X, Su S, Zhu S (2018) Hierarchical complementary attention network for predicting stock price movements with news. *ACM Press, Torino*, pp 1603–1606
55. Ling W, Trancoso I, Dyer C, Black AW (2015) Character based neural machine translation. [arXiv:151104586\[cs\]](https://arxiv.org/abs/151104586)
56. Golub D, He X (2016) Character level question answering with attention. [arXiv:160400727\[cs\]](https://arxiv.org/abs/160400727)
57. Ding C, Cheng HK, Duan Y, Jin Y (2017) The power of the “like” button: the impact of social media on box office. *Decis Support Syst* 94:77–84. <https://doi.org/10.1016/j.dss.2016.11.002>
58. Aizawa A (2003) An information theoretic perspective of tf idf measures. *Inf Process Manag* 39:45–65
59. Kratzwald B, Ilić S, Kraus M et al (2018) Deep learning for affective computing: text based emotion recognition in decision support. *Decis Support Syst* 115:24–35. <https://doi.org/10.1016/j.dss.2018.09.002>
60. Cho K, van Merriënboer B, Gulcehre C et al (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. [arXiv:1406.1078\[cs,stat\]](https://arxiv.org/abs/1406.1078)
61. Vaswani A, Shazeer N, Parmar N et al (2017) Attention is all you need. [arXiv:1706.03762\[cs\]](https://arxiv.org/abs/1706.03762)
62. Mnih V, Heess N, Graves A, Kavukcuoglu K (2014) Recurrent models of visual attention. [arXiv:1406.6247\[cs,stat\]](https://arxiv.org/abs/1406.6247)
63. Mikolov T, Chen K, Corrado G, Dean J (2013) Efficient estimation of word representations in vector space. [arXiv:1301.3781\[cs\]](https://arxiv.org/abs/1301.3781)



- 
64. Pennington J, Socher R, Manning C (2014) Glove: global vectors for word representation. Association for Computational Linguistics, Doha, pp 1532–1543
  65. Kim Y (2014) Convolutional neural networks for sentence classification. [arXiv:1408.5882\[cs\]](https://arxiv.org/abs/1408.5882)
  66. Reitsma JB, Glas AS, Rutjes AWS et al (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 58:982–990. <https://doi.org/10.1016/j.jclinepi.2005.02.022>
  67. Hung C, Lin H K (2013) Using objective words in SentiWordNet to improve word of mouth sentiment classification. *IEEE Intell Syst* 28:47–54. <https://doi.org/10.1109/MIS.2013.1>
  68. Li J, Chen X, Hovy E, Jurafsky D (2015) Visualizing and understanding neural models in NLP. [arXiv:1506.01066\[cs\]](https://arxiv.org/abs/1506.01066)
  69. Deng W, Wang Y, Liu Z et al (2014) HemI: a toolkit for illustrating heatmaps. *PLoS ONE* 9:e111988. <https://doi.org/10.1371/journal.pone.0111988>
  70. Wu C, Wei Y, Chu X et al (2018) Hierarchical attention based multimodal fusion for video captioning. *Neurocomputing* 315:362–370. <https://doi.org/10.1016/j.neucom.2018.07.029>
  71. Devlin J, Chang M W, Lee K, Toutanova K (2018) BERT: pre-training of deep bidirectional transformers for language understanding. [arXiv:1810.04805\[cs\]](https://arxiv.org/abs/1810.04805)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.