WILEY | Hindawi

Research Article

Data Mining Algorithm for Demand Forecast Analysis on Flash Sales Platform

Mingyang Zhang [iD],[1] **Yixin Wang,**[1] **and Zhiguo Wu** [iD][2]

[1]*Department of Management Science and Engineering, School of Economics and Management, Beijing Forestry University, Beijing 100083, China*
[2]*Department of Logistics Management, School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Mingyang Zhang; mingyangzhang@bjfu.edu.cn

With the development of the digital economy, the emerging marketing strategy of the e-commerce flash sales has been changing the traditional purchasing habits of customers. This imposes new decision-making challenges for companies involved in flash sales. It is important for companies to build the accurate product demand forecast analysis focusing on the characteristics of the flash sales and customer behaviors. In this paper, VIPS (Weipinhui, a Chinese e-commerce platform) is taken as a case study with the key focus on how sentiment factors in customer reviews affect product demand in flash sale platforms. The paper adopts two sentiment analysis methods based on emotional dictionaries. The method with a higher evaluation index is adopted to integrate the emotional factors into the autoregressive model for product demand and assessment. The experiments prove that the autoregressive model for integrating the sentiment factors demonstrates better forecasting performances than the models without sentiment factors. The experiments further confirm that when product demand for the previous two weeks and customer review sentiment factors in the previous week are taken into consideration, demand forecast effects are most accurate.

1. Introduction

With the outbreak of COVID-19 in 2020, a new round of industrial revolution has emerged in human society. As the number of customers in e-commerce platforms keep surging, traditional customer habits have also been changing accordingly. Flash sales, originated from the French sales platform VentePrivce.com, is a special sales platform featuring the business-to-customer (B2C) e-commerce platform. The website regularly launches all sorts of famous retail products and sells them at a relatively low discounted price to the website members. Compared to normal online shopping, its strong features such as limited shopping time, quantity, and low prices are more appealing to customers [1, 2]. VIPS (VIPS is an acronym for Vipshop Information Technology Ltd.) is a pioneer in the domestic e-commerce platform sector. VIPS launched the flash sales model and is one of the most successful in operation. Since its establishment in August 2008, VIPS has been adopting the flash sales business model featuring brand discounts, flash sales, and authenticity guarantee. Many countries' e-commerce platforms have used FS platforms, e.g., HauteLook.com, Limango, and Giltand. Brands negotiate a time period with the platform for flash sales, where they sell products at a lower price and restore the original price outside of the flash sale time. For customers, such a business model mainly demonstrates the fun of panic buying as the products are sold in limited quantities at low prices. When customers secure products, they crave while others fail and they will enjoy a sense of achievement and satisfaction. So far, VIPS boasts a total of 340 million members with strong loyalty and stickiness with a repurchase rate of above 87%. Since its conception, VIPS has only spent three years in getting successfully listed on the New York Stock Exchange. Up to

now, VIPS has more than 30,000 brand suppliers in partnership. It is number 1 with 38.1% of the market share in the whole Chinese online special sales market. By the third quarter of 2019, VIPS has registered 28 consecutive profit-making quarterly performances and broken the industry record. The key to VIPS' success lies in its adoption of the flash sales model (VIPS. About us. https://www.vip.com/about-us).

Traditional flash sales websites (such as Gilt and Rue La La.) do not have sections for customer reviews. In 2019, VIPS has modified its website modules and added the customer review module. The online word-of-mouth effects greatly impact customers' decision-making, and customer reviews stand as one of the most important formats of communication [3]. Before online shoppers make their decisions, they usually resort to customer reviews and evaluations, as well as horizontal comparison to identify product information on quality so as to reduce uncertainty [4]. Therefore, reviews are regarded as one of the key drivers to future product demand [5]. Analysis of the sentimental factors in customer reviews is an important method to understand customer thinking. We take the flash sales platform of VIPS as an example to analyze how customer reviews impact flash sales' product demand.

To date, research on flash sales e-commerce is relatively limited. The literature on the flash sales e-commerce model mainly adopts the research perspective from the customer's, the retailer's, and the platform's side. The literature is predominantly focused on the model's impact on customers' decision-making and on the exploration of the psychological mechanism of the purchase urge. Also, there are very few empirical sentiment analysis studies for flash sale patterns. In addition, the research of inventory management systems mainly focuses on optimizing models and algorithms, considering factors such as logistics and location, and less on considering emotional factors [6]. Using mathematical and optimization methods, the existence of the optimal solutions is proved, and then a simple heuristic algorithm is presented to maximize total inventory profit and determine the best values of variables [7] build a systematic and improved optimization model of the supply chain inventory. They proposed ant colony algorithms and fuzzy modelling [8] develop the open-source software JSOptimizer that can be used to optimize simulation models of complex engineering systems built with JaamSim, and solve several instances of the optimization problem. Zhang et al. [9] applies a matching model of inventory control strategy for material classification in practice, and demonstrates the applicability and feasibility of the model. This paper is meant to help companies draft more accurate inventories, and restocking plans before the flash sales kickoff to prevent customer loss due to shortage of supply and to prevent inventory fee increases due to overstocking.

The research framework for this paper consists of five parts: chapter one offers an introduction to the research background and puts forward the research topic. Chapter two presents the literature review and states the foundation and bases of the research. Chapter three describes the methods and processes for reviewing data collection, states the data reviewing preprocessing method, and introduces the method of word segmentation in the end. Chapter four establishes the sentiment analysis model based on the short-term forecast. A comparison of two dictionary-based sentiment analysis algorithms are made to select the one with the better evaluation index and integrate it into the model as the sentiment factor. In the end, the short-term forecast model is established for making forecasts and conducting experiments and analyses. Chapter five summarizes the major research conclusions, makes recommendations to the platform managers and developers from the perspective of operation management, and proposes directions for future improvement for deficiencies in the paper.

2. Literature Review

Currently, there is still a relatively small number of academic researches into the flash online sales business [10]. Discovers through the empirical study that the flash sales model can further stimulate customers' desire to purchase. The less educated customers are more likely to believe in the handsome amount saved, which will drive them to place the order. A pricing strategy for factories to offer a time-bound discount which will later expire will create more benefits. Huang and Benyoucef [11] believe that the flash sales e-commerce model is beneficial for establishing brand loyalty, increasing sales, and advancing the speed of destocking. Peng et al. [12] state that the perceived value of a product is based on three criteria: the function, the emotion, and the social interaction. They discover that the perceived value has a positive correlation to purchase willingness. Time pressure works on the perspectives of perceived values emotionally and socially to create negative impacts on purchase willingness. Ferreira et al. [13] forecasts a product's future demand through machine learning and optimizes product prices on the flash sales platform. Zhang et al. [1] identified that the expectations during the sales stage depend on the reputation effect, the price for flash sales, and the inventories for flash sales.

The characteristics of a flash sales platform, such as a discount, a quantity restriction, and a time restriction are similar to the daily deals or deal-of-the-day promotions or retail outlet stores. But there are some differences between them.

(1) The first difference is the source of the product; on the flash sales platform, customers can buy products from various regions, even various countries. Groupon is a famous daily deal website. Gao and Chen [14] said "These online voucher vendors sell vouchers in specific cities at discounts ranging from 50% to 90%. These vouchers are typically offered by local businesses, such as restaurants and spas." Krasnova et al [15] mentioned that "Deal-of-the-Day (DoD) platforms have quickly become popular by offering savings on local services, products, and vacations."

(2) The deal-of-the-day or daily deal more emphasizes economies of scale; itis different from flash sales [16]. Provided that "the deal-of-the-day (or daily deal) is a group-buying website, where buyers with similar purchase interests congregate online to obtain group discounts. For interested buyers to enjoy the daily deal, the number of confirmed buyers on the particular day has to exceed the minimum required number as indicated on each website." On the FS platform, the number of confirmed buyers has no restriction.

(3) Like the daily deal website Groupon, the bricks-and-mortar shop knows how many voucher was sold; it equals the shop know a part of demand in advance. But on the FS platform, the sold quantities in the FS period are not equal to the part demand of the ricks-and-mortar shop; it will affect only the demand of the shop.

In conclusion, the literature on the flash sales e-commerce model mainly adopts the research perspective from the customer's, the retailer's, and the platform's side. The literature is predominantly focused on the model's impact on customers' decision-making and on the exploration of the psychological mechanism of the purchase urge. Despite the increasing popularity of the FS in practice, the number of literature papers on the impact of customer reviews on customer decision-making in flash sales platforms is relatively few. This paper is developed on the foundation of previous reviews. It adopts the perspective of flash sales platforms, carries out empirical studies based on real sales numbers and review data, and aims at enriching the research content of demand forecast in flash sales e-commerce models.

At the current stage, text sentiment analysis is an important research branch in the field of web data mining. It is widely used in real life. Other widely used models include: classification models [17], recommendation system [18], customer relationship management models [19], stock market prediction [20], social problems monitoring [21, 22], opinion polling [23], and competitive intelligence acquisition [24]. Sentiment analysis technologies mainly include the machine learning method and the semantic orientation method. The machine learning sentiment categorization method requires a large amount of sample training in application to set up [25]. Makes good use of the N-gram words and their special features. For the first time, they apply naïve bayes, support vector, and maximum entropy into passage-level sentiment categorization tasks. The semantic orientation method focuses on the subtraction of sentiment words and judgment of the sentiment polarity. Therefore, it does not require training beforehand [26]. Stacked denoising autoencoders (SDAs) were used to provide an infrastructure to resolve issues of sentiment recognition from textual contents. The results indicate the promising capability of SDAs to perform sentiment recognition on a multitude of domains and languages [27]. Proposes an improved stacking framework which contains multiple layers for predicting whether the stock price index will increase or decrease with respect to the price prevailing sometime earlier, if necessary, a month [28]. Build a domain-dependent sentiment dictionary, SentiDomain. They propose a weak supervised neural model that aims to learn a set of sentiment cluster embeddings from sentence global representation of the target domain. Kumar et al. [29] propose an efficient method for sentiment analysis by using particle swarm optimization, which experiments show that the proposed technique outperforms other state-of-the-art techniques. Hu and Liu [30] judge the sentiment polarity for words selected from dictionaries and complete the categorization by calculating the weighted sum.
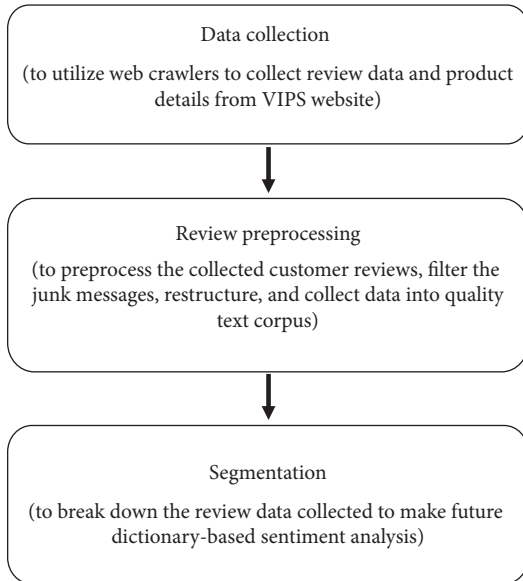
Up till now, barely any flash sales platforms have launched with the customer review block. The paper will make use of the customer review information in VIPS and apply the sentiment analysis method to produce a better demand forecast. The demand forecast model is usually realized through an autoregressive model (Water 2004), linear regression model [31], time series association analysis (Chatfield 1984), Granger causality analysis [32], and nonlinear model optimization [33]. Using the blog sentiment analysis to forecast box office sales [34], the autoregressive emotion sensitive model (ARES) was proposed. They discover that the sentiment captured in the blog one day before has the best forecasting results for the box office sales prediction for the next day [35]. It is discovered that when the film is released, heated discussion will take place in the microblog platform. Later, the number of relevant blogs will gradually decrease. The box office returns go through similar processes. Return on the Sunday of the release week is usually the highest. Therefore, they construct a forecast model based on linear regression to forecast box office return for the week of release [36]. A linear regression empirical study was carried out for audio, video plays as well as electronic cameras in Amazon [31]. Forecast transmission of infectious diseases by setting up a logistic regression model. Using blog mention count to forecast peak sales for books [37], we use a series of predictive classifiers such as Light GBM, XGBoost, Logistic Regression, and Random Forest in order to evaluate the probability of a customer entering loan default. Gruh et al. [38] utilize time series correlation analysis to forecast the timing advance. They found that different books in the samples have different timing advances. This may have to do with the fact that the arrival of a book's peak sales is subject to the occurrence of various social events. Forecasting stock market development using blog sentiment analysis [39], the Granger causality analysis was applied to analyze various sentiment time series and the Dow Jones index time series. They believe that the best timing advance for the forecast is two days [40]. The impacts of product reviews in a competitive market were proved [41]. It is identified that the sentiment value in product reviews has a significant impact over future product demand [42]. The convention rate of the influence was studied from sentiment value in online reviews.

The above-mentioned research shows that the autoregressive model, as a practical model for addressing the problem of time series, has already been widely applied to various forecast scenarios. A short-term sentiment aware

autoregressive model (SAAR) can be established based on the sentiment factors and previous sales. The paper mainly investigates sentiment analysis methods to study social behavior, emotional dictionaries for flash sales patterns. Authentic product sales numbers and review data are adopted to verify the model and guarantee authenticity and accuracy of the research.

3. Data Collection and Data Processing

There are three steps for data collection and data processing:

```
┌─────────────────────────────────────────────┐
│           Data collection                    │
│ (to utilize web crawlers to collect review   │
│  data and product details from VIPS website) │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│          Review preprocessing                │
│ (to preprocess the collected customer        │
│  reviews, filter the junk messages,          │
│  restructure, and collect data into quality  │
│  text corpus)                                │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│             Segmentation                     │
│ (to break down the review data collected to  │
│  make future dictionary-based sentiment      │
│  analysis)                                   │
└─────────────────────────────────────────────┘
```

The paper uses GooSeeker as the tool for data collection and selects the domestic appliances categories in the VIPS platform for data mining. In every category, 2 brands from 17 types of products will be reviewed (in order to clearly figure out what impact of consumer reviews on product's demand, we only consider the two brander case). In regards to reviewing data, the collection is done in a reverse manner. In other words, data are collected from the very day on and backward. For each one of the products, data mining will end when the date reaches December 30th, 2019. In total, 10,000 reviews have been collected. In terms of the product details, the paper has collected the real-time domestic appliances popularity ranking list from December 30th, 2019, to March 29th, 2020. Product details are collected on a regular daily basis, which has amounted to 1547 items. The data mining rules in GooSeeker's MS station have been adopted and customized into those of VIPS' own. For each review, the time, customer name, customer level, and review text have all been noted.

As many low-quality items in text reviews may affect future analyses, five steps need to be taken in the preprocessing stage as follows:

(1) Removing duplication: removing duplication is to delete repetitive messages in the customer review. The same person may purchase multiple times in one store, which leads to repetitive reviews. In such circumstances, only the earliest review will be saved and the remaining repetitive ones will be deleted.

(2) Mechanical compression: this step is to process repetitive parts in the sentence. In the paper, the redundant part of the text is processed, mainly centering on the beginning and the end of sentences. For instance, in "Thumbs up. Not baaaaaaaaaaaaaaaaaad." only "bad" needs to be reserved. Otherwise, the future sentiment value calculation will not be affected. As a result, the whole sentence is compressed into "Thumbs up. Not bad."

(3) Short sentence compression: short sentence compression is mainly about deleting extremely short or meaningless reviews. In this paper, texts with less than five international characters are deleted. Short sentences include sentences that are short to start with and those that become short after mechanical compression, i.e., those long texts featuring meaningless repetitiveness.

(4) Removing emojis and emoticons: the last step is mainly about manually deleting emojis and emoticons in sentences.

(5) For Chinese sentences: segmentation will be imposed on the word series. After segmentation, the sentence 'I am very satisfied. Looking Good. I'm so into it.' will be turned into '/I'm/very/satisfied/, Looking/Good. I'm/so/very/into/it.' After removing stop-words, the review will look like 'Very/satisfied. Not bad/So/into it.'

In order to judge whether there are sentiment words contained, we need to segment every review and accurately keep the keywords. Accuracy of the segmentation means a lot to the following analysis. Therefore, the method with better effects needs to be selected. For Chinese sentence segmentation, many methods are available. This paper has resorted to '*Jieba*' and Python's Chinese segmentation package and to handle reviews in text documents. Accuracy, efficiency, generality, and applicability are the most important factors in segmentation performances. This system offers more than 97% accuracy [43] and features easy installation, extensive language support, and quite a degree of popularity. After segmentation, unnecessary words need to be removed. Unnecessary words include prepositions, pronouns, function words, and characters irrelevant to sentiment analysis. After preprocessing, reviews for the 17 products will be segmented.

Thanks to '*Jieba*' above 97% accuracy [43], the approach is used in this paper to segment the texts. The paper adopts the generative model based on the *Snownlp* stop-word list (https://github.com/isnowfy/snownlp). The negative words and degree level adverbs are filtered to generate a new stop-word list.

4. Sentiment Aware Model (SAM) for Short-Term Forecast

4.1. Model Assumption. First of all, the relationship between sentiment value and product demand in product reviews is investigated. The experimental data are used to test such a correlation. Then, the model assumption is proposed.

The paper takes the domestic appliances in the VIPS flash sales platform as the target of research and collects customer reviews during a designated period. However, as there is no direct access to the sales figures, the number of reviews published can be roughly taken as the demand for the domestic appliances.

Here it is assumed as follows:

H1: the number of reviews for a product equals the demand for the corresponding product.

Due to the unique nature of the flash sale platform, we cannot see the sales quantities of the product on VIPS, and the platform does not show bad reviews, only positive ratings and the number of product reviews. Moreover, Park et al. [44].found that purchasing intention increased as the number of reviews increases. Therefore, we made the assumption of H1.

Product review mining is an important application of sentiment analysis. Scholars adopt various econometric models and research methods to measure the enterprise communication effects of online product reputation in multiple dimensions. The three most commonly used dimensions are volume, valence, and dispersion. Volume is mainly referring to the number of customer reviews for a certain product. It reflects the awareness effect of the online reputation [45]. According to the rules of VIPS, reviews will only be shown when the total number reaches 999. Therefore, there is no direct access to the total review number. That's also why the impact of volume on sales is not considered for now. Dispersion means the degree of communication in different online communities. The higher the dispersion is, the greater the influence. Because this paper only looks at the VIPS community, measuring dispersion is not applicable. Therefore, valence is the major dimension used to analyze the impact of online reputation on product demand.

Valence measures the customers' feedback on the products in both good and bad, positive and negative ways and is usually measured with an overall score (good-bad) or a ratio between the good and bad (good/bad). It reflects the persuasive effect of a product's online reputation [46]. It is discovered that improvement of book reviews can increase demand for the book. In the meanwhile, drops in demand caused by negative reviews are more prominent compared to the increase in demand incurred by positive reviews. Similarly, Floh et al. [47] found that the stronger a review is, the more likely it will stimulate purchase increases or decreases. In other words, intense positive or negative reviews create greater influence than those with mixed emotions.

Based on the analysis above, the following assumption is made:

TABLE 1: Sale level and comments emotion score.

	Sale level	c
Week 1	70	8.58
Week 2	57	9.21
Week 3	63	9.77
Week 4	34	6.59
Week 5	38	6.67
Week 6	69	6.59
Week 7	67	11.40
Week 8	117	17.66
Week 9	109	18.15
Week 10	112	10.59
Week 11	93	9.36
Week 12	78	10.22
Week 13	87	7.35

H2: a reviewer's overall emotion (demonstrated through written feedback) toward the product creates a positive or negative impact on product demand.

The demand data have been acquired for "Media 304 Stainless Domestic Electric Kettle 1512d" in 13 weeks, in 944 review messages, through the web crawler technology. Using the improved SAM, the sentiment value in product reviews is calculated as shown in Table 1.

The SAAR model proposed by Liu et al. [34] reveals that the sentiment information captured in blogs can achieve the best performance in forecasting the film box office ticket sales for the next day. Using the number of mentions in blogs to predict books' peak sales volumes [38], the time series correlation analysis was utilized to confirm the advancement time for the prediction. They identify differences between books. However, basically, the gap ranges from several days to several weeks. Therefore, the time lag is uncertain, and the advancement is usually affected by the scope of application, social behavior, and the method of sentiment analysis. The current sentiment value may have an impact on the product demand for the next cycle. When the correlation analysis is carried out, the impact from the sentiment value is delayed. If it is assumed that the time lag is one cycle [34], then that means sentiment value in the first week affects demand in the second week, and so on so forth. Therefore, the data to be analyzed for correlation is $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n))$, and the formula for calculating the ratio r is shown in the following formula:

$$r = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 \cdot \sum_{i=1}^{n} (y_i - \overline{y})^2}}. \tag{1}$$

In the formula, x_i is the demand for week i and $\overline{x}$ is the average demand within the observation period. y_i is the sentiment value in week i and $\overline{y}$ is the average sentiment value within the observation period.

The calculation results are shown in Table 2.

Such experiment results to demonstrate a strong correlation between the sentiment value and demand at the 0.01 level (two sides).

Based on the analysis above, the following assumption is made:

Table 2: Correlation analysis results.

		Demand	Sentiment value
Demand	Pearson correlation	1	0.721**
	Significance (double tail)		0.008
	Number of cases	12	12
Sentiment value	Pearson correlation	0.721**	1
	Significance (double tail)	0.008	
	Number of cases	12	12

Table 3: Example of sentence sentiment score.

Vocabulary	Sentiment score
Can't turn around (hui tian Wu li)	−4.74492768973
Speechless (wuyu)	−3.66149641358
Anxious (ganzhao ji)	−2.17885364262
No second to (buyayu)	0.858219035923
Bravo (piao liang)	2.23552351188
Super value (chao zhi)	3.44664867894

Table 4: Degree level word weights.

Degree level words	Example	Weight
"Most"	One hundred percent and extremely	2
"Pretty"	Pretty and very much	1.5
"Quite"	Quite	1.25
"Rather"	Rather	0.5
"A little"	Not quite and a little	0.25
"Too"	Too and over	0.1

H3: current sentiment value impacts demand during the next cycle.

4.2. Dictionary-Based Sentiment Analysis Method. Dictionaries for emotions and degree level words have been compiled. First of all, the *Boson NLP* emotional dictionary is selected to judge the sentiment scores. Negative scores represent more negative words. Positive scores represent more positive words. The degree of emotions can be reflected through the scores. Table 3 shows examples of words and their sentiment scores based on the *Boson NLP* emotional dictionary.

This paper uses a degree level adverb dictionary and the integrated negative word dictionary from the sentiment analysis vocabulary (beta version) in the cnki network. Customers often resort to degree level adverbs and negative adverbs in expressing emotions. For instance, they may use degree level adverbs (e.g., 'quite', 'extremely', 'somewhat', 'a little') to emphasize the subtle differences in emotions. Also, some negative adverbs like "not" will change the sentiment polarity. The sentence "she is not beautiful." is an example. The degree-level word list is shown in Table 4. Based on the relevant word information provided by cnki, a certain weight is given to the common degree level adverbs in the corpus. According to cnki, the degrees are categorized into six levels: extremely (most), pretty, quite, rather, a little, and too. The weight given in this paper is noted as follows:

W is set as the weight of the degree level word and S as the sentiment word value. Sentiment_index is the subscript for the sentiment word. Calculation of the sentiment degree is as follows:

According to the above Algorithm 1, for the example of "I'm satisfied. The look is pretty good." The calculation of the sentiment in the sentence works as follows:

(1) The words 'satisfied', 'pretty', and 'good' are left after preprocessing the data and removing stop-words like 'I'm', 'look', and 'is.'

(2) The sentiment words 'satisfied' and 'good' have weights of 2.84 and 2.65. The degree level word 'pretty' has a degree value at 1.52.

(3) Because there is no negative word, the negative sentiment value is 0.

(4) Therefore, the sentiment value of the sentence is $2.84 + 2.65 * 1.52 = 6.868$.

In the end, the experiment effectiveness is evaluated. Because the sentiment analysis software for *ROSTCM6* is based on optimizing emotional dictionaries, the accuracy is higher than those based on word vectors or neural networks. This paper uses *ROSTCM6* in analyzing the sentiment value in customer reviews. For *ROSTCM6* software, the experiment results include positive emotion, negative emotion, and neutral emotion. In this paper, if the sentiment value of the sentence is greater than 0, then it means positive emotion. If the sentiment value is less than 0, then it means negative emotion. The benchmark is marked out manually based on the condition that it is all correct and does not involve individual differences. The closer the result is to the benchmark, the more accurate the model proves to be.

Three major assessment indexes are adopted here: recall, precision, and *F*-measure.

(1) *Recall rate*: investigating the comprehensiveness of the sentiment categorization model and reflecting the ratio of the number of correctly identified to the number of identified total after the experiment.

$$\text{Recall rate} = \frac{\text{correctly identified ones}}{\text{identified total}}. \quad (2)$$

(2) *Precision rate*: investigating accuracy of the model and reflecting the ratio between the number of the correctly identified against the number that ought to be identified after the experiment.

$$\text{Precision rate} = \frac{\text{correctly identified ones}}{\text{actual total}}. \quad (3)$$

(3) *F-measure*: the harmonic mean of the two when the recall rate and the precision rate are viewed as equals.

$$F - \text{measure} = \frac{(2 * \text{recall rate} * \text{precision rate})}{(\text{recall rate} + \text{precision rate})}. \quad (4)$$

The paper experiments with 100 reviews for the first product. From results in Table 5, it can be noted that the tree

```
(1) for screening the segmentation result, do
(2)    if the word belongs to the sentiment vocabulary, then
(3) score+ = W * S.
(4) sentiment_index+ = 1
(5)    if sentiment_index is smaller than the total amount of all sentiment word, then
(6) for degree level adverbs or negative words that exist between the current sentiment word and the next, do
(7) if it is a negative word, then
(8) W* = −1.
(9) end if
(10) if there is any degree level word, then
(11) W* = V
(12) end if
(13) end for
(14) end if
(15) end if
(16) end for
```

ALGORITHM 1

assessment indexes in the new model are all higher than those in the original model. The optimized model shows significant performance improvement. The paper adopts the current sentiment analysis model to develop SAM based on short-term forecasts.

4.3. Model Development. Sales data are used instead of demand data to predict current product demand using previous sales performances. In real life, the current sales of a product show a certain correlation with its previous sales. Therefore, the autoregressive model is more suitable. The domestic appliances in VIPS are our research object. Customer reviews for the domestic appliances in VIPS are collected for a certain period. As there is no direct access to the product sales numbers, the number of reviews for a product is taken as the approximate number of product sales. Affected by multiple factors such as the seller preparation time, the delivery speed, postponement of buyers' feedback, and looking at the data on a daily basis, there may be days when there are zero reviews or a huge amount of reviews. Data with such big fluctuations are apparently not applicable for model development. In order to reduce impacts from fluctuations, we take weeks as the time series unit. The demand for the product over 13 weeks in VIPS is captured as follows:

According to the autoregressive distributed lag, it is forecasted that the demand for the domestic appliance during the time-frame X_t requires p periods before X_t remains stable. Otherwise, different treatment will be carried out. As in Figure 1, the demand features prominent fluctuations during different phases and different treatment is necessary. By calculating the logarithm for each element in the demand series $\{x_t\}$, a new demand series $\{y_t\}$ is produced. Please refer the following formula:
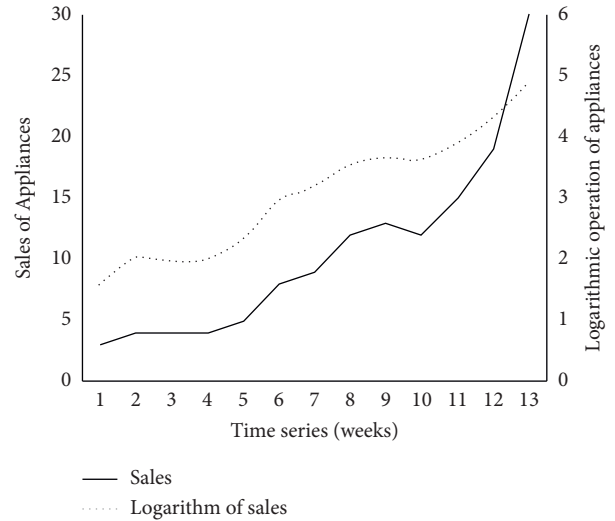
$$y_t = \log_2(x_t). \tag{5}$$

FIGURE 1: Example of domestic appliance sales

After processing the data, the autoregressive coefficient must be forecasted, and then, the observed data must be fit into a linear parameter model. In this paper, the estimate will be carried out upon the training group by ordinary least square. In the end, the model is used to analyze and investigate relations between demand data in different phases to forecast current demand.

To apply the autoregressive distributed lag (ADL) model, ADF and other relevant tests need to be passed. To secure the stability of the data, an ADF test needs to be applied to the new demand series. In this paper, the critical value of 0.05 is taken as the standard. If the data fails the test, they will be directly abandoned. After the ADF test, 12 categories of domestic appliances pass. The autocorrelation (AC) coefficient is to be calculated for the new demand series $\{y_t\}$ after the ADF test. AC coefficient between variables is needed for

the model. If the coefficient is greater than 0.5, the AC model should be set up. Otherwise, the data should be abandoned. After the test of the AR coefficient, 10 domestic appliances survive.

Besides previous sales performance of the product, customer opinions also have influence on the current sales. Therefore, the sentiment factor is brought in to optimize the model. If we take C_t as the number of product reviews during the observation period t and set the observation cycle at one week, the average sentiment value during t period is defined as in the following formula:

$$S_t = \frac{\sum_{c=C_t} e}{|C_t|}, \tag{6}$$

S_t is the average sentiment value during the observation period t. e is the value after *ROSTCM6* is used to generate sentiment analysis. S_t will be integrated with the autoregressive model to obtain a short-term-based SAM as shown in formula (7). In essence, it is an application of the ADL model.

$$y_t = \sum_{i=1}^{p} \theta_i y_{t-i} + \sum_{j=1}^{q} \lambda_j S_{t-j} + \varepsilon_t. \tag{7}$$

y_t is the product sales as a function of time t. S_t is a sentiment element function of t. q and p are parameters selected by the users. Parameter q is selected by users, and it is sentiment information from a few weeks ago, while p is sales information from a few weeks ago. θ_i is the demand coefficient in history, λ_j is the sentiment coefficient, and ε_t is the error term (white noise with an average value of 0).

4.4. The Model Experiment Results and Analysis. The paper carries out the ADF test and autocorrelative examination after data processing. It selects demand data that can be used, which is data for ten domestic appliances in VIPS over 13 weeks and categorizes them into the training group and test group. In the training group, the study is carried out toward the coefficient θ_i $(i = 1, 2, \ldots, p)$ and λ_i $(i = 1, 2, \ldots, q)$ in the model by ordinary least squares.

The paper evaluates the model effectiveness with mean absolute percentage error, MAPE. The calculation of MAPE is shown in the following formula:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{n} \frac{\left| \text{tred}_i - \text{true}_i \right|}{\text{true}_i}. \tag{8}$$

In the formula i is week number, $(i = 12, 13, \ldots, n)$, n is the estimated time in total, Pred_i is the estimated value obtained using the model. True_i is the actual value. The smaller the MAPE is, the better the model is in making forecasts.

The paper conducts the ADF test using *EVIEWS10*. Only those that withstand the test can be applied to the

autoregressive model with the parameters filled in place to generate the forecast value. After processing, the MAPE value can be generated. The experiment is carried out on the condition of $p \in [1, 6]$ and $q \in [1, 6]$. The combination of p and q with the best result is selected. The data for the first 11 weeks are taken as the training group, the data for the last two weeks are taken as the test group, and then a forecast is made for every domestic appliance in the two weeks test period. First, parameter q in the optimized model is fixed at 1 and tests are carried out on the condition that $p \in [1, 6]$. Results are shown in Figure 2.

As is shown in Table 6, when p equals 6 or 5, the Sam outperforms the autoregressive model. When p equals 1, 2, 3, and 4, the forecast effect is even more prominent. It shows certain relations between the forecast effect and the sentiment factors.

When p equals 5, the effect of the SAM is not easily observed. When p equals 6, the autoregressive model has better performances than the SAM. When p equals 5 and 6, the average value of MAPE in the SAM is bigger than that in the autoregressive model. This might be a result of influence from sentiment words in the reviews. In the autoregressive model, the MAPE values for products at 2, 7, and 8 are about 7% lower than those in the independent AR model. Upon observation of the data, neutral emotions take up 1/3 of the total, accounting for quite a prominent percentage. The major reason is that within a certain cycle t, and the neutral emotions contained in sentiment values in product reviews only increases the number of reviews. It is for sure that the ultimate sentiment information will be weakened. However, in real life, the neutral reviews often contain relatively complicated messages, details of which cannot be easily processed, and generate biased results.

When p equals 2, effect of parameter q on the model is shown in Figure 3. The result shows the worst performance of the model when q equals 1. That means that the sentiment information in reviews published one week before has the best effect for forecasting demand.

Figure 4 shows the overall situation for ten different electronic products. It can be noted from the figure that, in the autoregressive model and the SAM, when $p = 2$ and $p = 6$, the best and worst cases, respectively, happen. It means the selection of p itself also affects accuracy of the model. If the value of p is too small, the hidden relationship between numbers can be easily ignored. However, when the value of p is too big, there will be too much distracting data. In the experiment, when $p = 2$, the best effect is achieved, which means sales of a certain week is affected by sales of the previous two weeks. It has to do with the time and frequency of flash sales. The research object of the paper has been arranged for flash sales promotion in weeks 6 and 7 and weeks 10 and 11. Sales performances in the week before flash sales and in the first week during flash sales both have impacts on sales in the second week during the flash sales period. That means the reputation of the product itself and
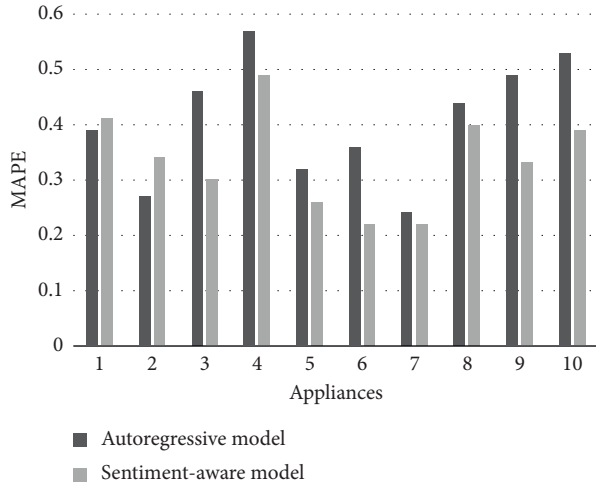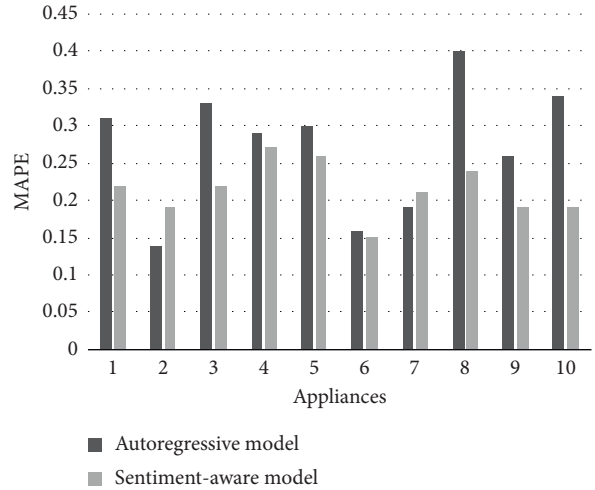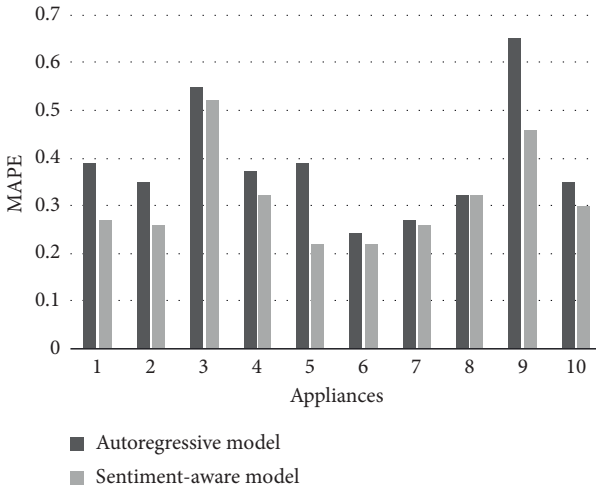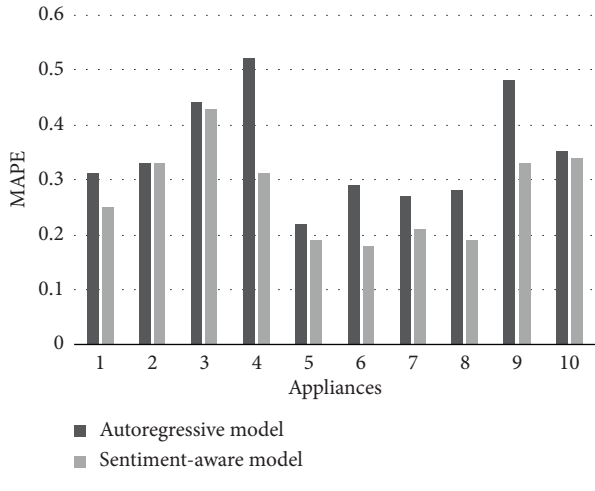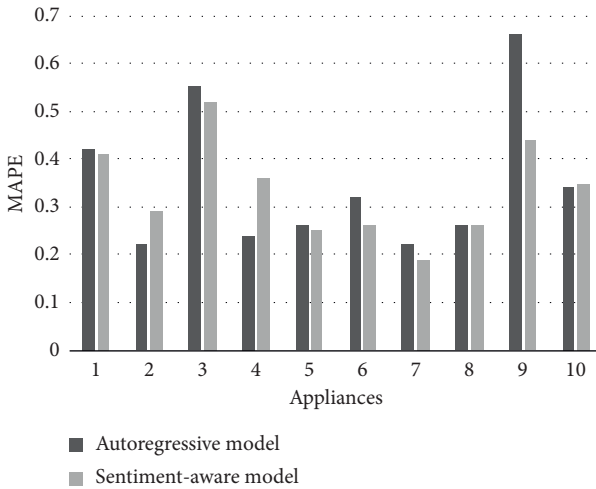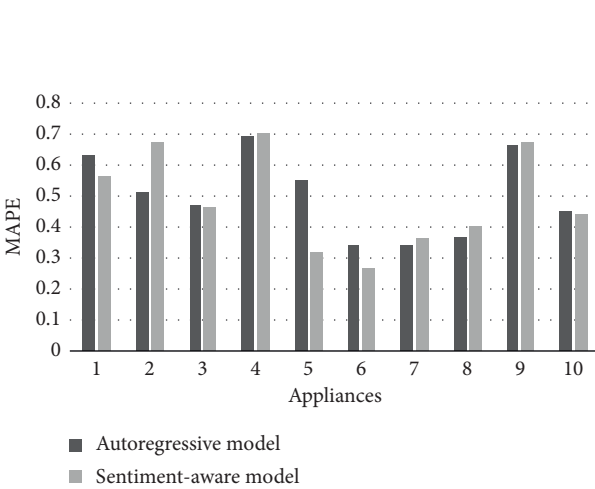
Figure 2: Experimental results of appliances sales forecast, (a) results of sales information one week ago, (b) results of sales information two weeks ago, (c) results of sales information three weeks ago, (d) results of sales information four weeks ago, (e) results of sales information five weeks ago, and (f) results of sales information six weeks ago.

the format of flash sales are both influencing factors. For products with a good reputation, when the flash sales season comes to the end, customers will feel the time pressure and get further stimulated to purchase. For products with a bad reputation, a joint effect can still be noted when the flash sales period is about to end. Furthermore, the sales performance in the first and second week during the flash sales period also affect sales performance during regular time.

The experiment result shows that sentiment information can affect demand. When the sentiment factors are considered in the autoregressive model, the forecast effects register a prominent improvement. In inventory management, the forecast can be utilized to predict the number of orders for the next cycle to control an enterprise's overall stock and purchase numbers and to reduce the inventory cost. Regarding sales management, the forecast model should be first used for prediction. If huge fluctuations are identified, it is highly likely that certain sales strategies have been applied to competitive products. The use of the model is beneficial for the company as it can adjust the product price based on the competitive prices in a timely manner.

5. Discussion and Conclusions

The paper takes the VIPS as the object of research and prioritizes investigating how sentiment factors in customer reviews affect demand forecasts for products in the flash sales platform. The contribution of the paper is mainly summarized in the following three aspects:

(1) Based on the authentic sales data and review data from the flash sales platform, explorations are made regarding influencing factors on customer behavior in flash sales platforms. Correlation between sentiment factors and demand is proved through solid experimental results.

(2) A science-based analysis framework is offered to enterprises when they establish sentiment-oriented analysis model for product sales. The paper has adopted theories and methods related to a sentiment analysis and implemented sentiment mining on customer review data. Considering the special words and factors that may affect the customers' reviews sentiment analysis results, text data have been converted to numerical data to improve the original sentiment analysis model and increase its accuracy.

(3) Investigations are carried out to use the flash sales model in the correct way to forecast demand and enhance enterprise performances, especially for inventory optimization. The experiment has proved that the autoregressive model, which integrates the sentiment factors' features, leads to better forecast. Furthermore, the autoregressive model has best performances in terms of demand forecast, driven by customers' sentiment factors, when the forecast is targeted at one or two weeks beforehand.

TABLE 5: Comparison of evaluation indexes.

	Original model	Optimized model
Recall (positive)	0.87	0.95
Precision (positive)	0.76	0.91
Precision (negative)	0.31	0.53
Precision (negative)	0.50	0.69
Recall	0.59	0.74
F (positive)	0.81	0.93
F (negative)	0.38	0.60

TABLE 6: Analysis of experiment results.

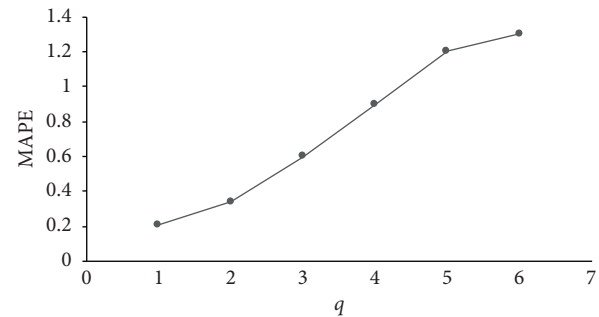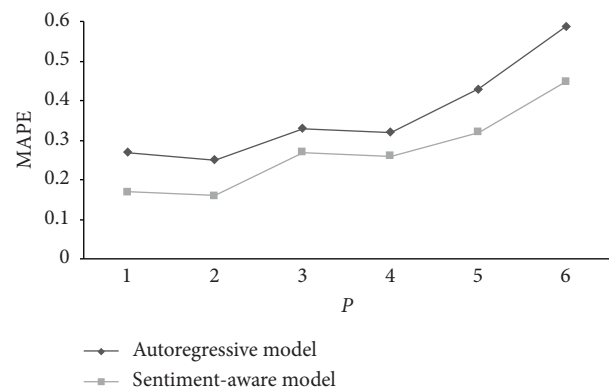| p	Percentage of SAM outperforming auto regressive model (%)	Average MAPE value in SAM against that in auto regressive model
1	80	0.85
2	80	0.85
3	90	0.83
4	90	0.80
5	60	1.01
6	50	0.97

FIGURE 3: Effect of parameter q on model.

FIGURE 4: Result of comparative experiment.

As for future research, the following is proposed:

(1) The Internet catchphrases and emojis have already been quite pervasive in nowadays reviews, which can also demonstrate the emotional attitude of the customers. The paper does not have dedicated research into emojis and only filters the relevant

information. In later stages, sentiment analysis can be carried out for both the trendy phrases and emojis.

(2) The quality of reviews is not taken into consideration, and the false reviews cannot be fully filtered. Therefore, the authenticity of the reviews can further be improved. In later stages, customer levels and thumbs-up numbers for reviews can also be considered to optimize the model.

(3) This paper adopts the time series data. However, the future study can resort to panel data and select multiple factors to establish the model for the empirical study of product demand.

Data Availability

The data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work. The authors declare that they do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgments

References

[1] M. Zhang, T. C. E. Cheng, and J. Du, "Advance selling of new products to strategic consumers on flash sale platforms," *International Journal of Logistics Research and Applications*, vol. 21, no. 3, pp. 318–331, 2018.

[2] M. Zhang, J. Zhang, T. C. E. Cheng, and G. Hua, "Why and how do branders sell new products on flash sale platforms?" *European Journal of Operational Research*, vol. 270, no. 1, pp. 337–351, 2018.

[3] B. Bickart and R. M. Schindler, "Internet forums as influential sources of consumer information," *Journal of Interactive Marketing*, vol. 15, no. 3, pp. 31–40, 2001.

[4] F. Zhu and X. Zhang, "Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics," *Journal of Marketing*, vol. 74, no. 2, pp. 133–148, 2010.

[5] N. Hu, L. Liu, and J. J. Zhang, "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects," *Information Technology and Management*, vol. 9, no. 3, pp. 201–214, 2008.

[6] R. Keshavarzfard, A. Makui, and R. Tavakkoli-Moghaddam, "A multi-product pricing and inventory model with production rate proportional to power demand rate," *Advances in Production Engineering & Management*, vol. 14, no. 1, pp. 112–124, 2019.

[7] W. Yu, G. Hou, and J. Li, "Supply chain joint inventory management and cost optimization based on ant colony algorithm and fuzzy model," *Tehničkivjesnik*, vol. 26, no. 6, pp. 1729–1737, 2019.

[8] D. Katsios, A. S. Xanthopoulos, D. E. Koulouriotis, and A. Kiatipis, "A simulation optimisation tool and its production/inventory control application," *International Journal of Simulation Modelling*, vol. 17, no. 2, pp. 257–270, 2018.

[9] Z. L. Zhang, Y. F. Wang, and Y. Li, "Inventory control model based on multi-attribute material classification: an integrated grey-rough set and probabilistic neural network approach," *Advances in Production Engineering and Management*, vol. 14, no. 1, pp. 1854–6250, 2019.

[10] D. Mattioli, "Macy's plan: boots, bieber," *The Wall Street Journal*, pp. 26-27, 2011.

[11] Z. Huang and M. Benyoucef, "User preferences of social features on social commerce websites: an empirical study," *Technological Forecasting and Social Change*, vol. 95, pp. 57–72, 2015.

[12] L. Peng, W. Zhang, X. Wang, and S. Liang, "Moderating effects of time pressure on the relationship between perceived value and purchase intention in social E-commerce sales promotion: considering the impact of product involvement," *Information & Management*, vol. 56, no. 2, pp. 317–328, 2019.

[13] K. J. Ferreira, B. H. A. Lee, and D. Simchi-Levi, "Analytics for an online retailer: demand forecasting and price optimization," *Manufacturing & Service Operations Management*, vol. 18, no. 1, pp. 69–88, 2016.

[14] F. Gao and J. Chen, "The role of discount vouchers in market with customer valuation uncertainty," *Production and Operations Management*, vol. 24, no. 4, pp. 665–679, 2015.

[15] H. Krasnova, N. F. Veltri, K. Spengler, and O. Günther, ""Deal of the day" platforms: what drives consumer loyalty?" *Business & Information Systems Engineering*, vol. 5, no. 3, pp. 165–177, 2013.

[16] Y. Liu and J. Sutanto, "Buyers' purchasing time and herd behavior on deal-of-the-day group-buying websites," *Electronic Markets*, vol. 22, no. 2, pp. 83–93, 2012.

[17] O. Voican, "Using data mining methods to solve classification problems in financial-banking institutions," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 54, no. 1, 2020.

[18] J. Yoon and S. Joung, "A big data based cosmetic recommendation algorithm," *Journal of System and Management Sciences*, vol. 10, no. 2, pp. 40–52, 2020.

[19] D. Kurunathan, S. Shanmugathas, and K. Ashoka, "Analysis of relation between customer behavior and information Technology market," *Journal of System and Management Sciences*, vol. 9, no. 1, pp. 87–104, 2019.

[20] Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Computing and Applications*, vol. 32, pp. 9713–9729, 2019.

[21] J. Awwalu, A. A. Bakar, and M. R. Yaakub, "Hybrid N-gram model using Naïve Bayes for classification of political sentiments on Twitter," *Neural Computing and Applications*, vol. 31, no. 12, pp. 9207–9220, 2019.

[22] L. Fu and Y. Dong, "Research on internet search data in China's social problems under the background of big data," *Journal of Logistics, Informatics and Service Science*, vol. 5, no. 2, pp. 55–67, 2018.

[23] M. Soleymani, D. Garcia, B. Jou, B. Schuller, S.-F. Chang, and M. Pantic, "A survey of multimodal sentiment analysis," *Image and Vision Computing*, vol. 65, no. 65, pp. 3–14, 2017.

[24] J. Roca-Gonzalez, J. Vera-Lopez, and G. Rodriguez-Bermudez, "Analysis of patent# US2014/0319274a1: a case study of simulations for new designs review," *International Journal of Simulation Modelling*, vol. 17, no. 3, pp. 405–418, 2018.

[25] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," arXiv preprint cs/0205070, 2002.

[26] H. Sagha, N. Cummins, and B. Schuller, "Stacked denoising autoencoders for sentiment analysis: a review," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 7, no. 5, Article ID e1212, 2017.

[27] M.-Q. Jiang, J.-P. Liu, and L. Zhang, "An improved stacking framework for predicting stock price index direction," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 53, pp. 187–202, 2019.

[28] M. Ahmed, Q. Chen, and Z. Li, "Constructing domain-dependent sentiment dictionary for sentiment analysis," *Neural Computing and Applications*, vol. 32, pp. 14719–14732, 2020.

[29] R. Kumar, H. S. Pannu, and A. K. Malhi, "Aspect-based sentiment analysis using deep networks and stochastic optimization," *Neural Computing and Application*, vol. 32, pp. 3221–3235, 2020.

[30] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168–177, Washington, DC, USA, August 2004.

[31] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," *Nature*, vol. 457, no. 7232, pp. 1012–1014, 2009.

[32] J. R. Freeman, "Granger causality and the times series analysis of political relationships," *American Journal of Political Science*, vol. 27, no. 2, pp. 327–358, 1983.

[33] Y. Yang and Z. Duan, "An effective co-evolutionary algorithm based on artificial bee colony and differential evolution for time series predicting optimization," *Complex and Intelligent Systems*, vol. 6, no. 4, 2020.

[34] Y. Liu, X. Huang, A. An, and X. Yu, "ARSA: a sentiment-aware model for predicting sales performance using blogs," in *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 607–614, Amsterdam, The Netherlands, July 2007.

[35] S. Asur and B. A. Huberman, "Predicting the future with social media," in *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, pp. 492–499, Toronto, Canada, August 2010.

[36] N. Archak, A. Ghose, and P. G. Ipeirotis, "Deriving the pricing power of product features by mining consumer reviews," *Management Science*, vol. 57, no. 8, pp. 1485–1509, 2011.

[37] A. Coşer, M. M. Maer-Matei, and C. Albu, "Predictive models for loan default risk assessment," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 53, no. 2, 2019.

[38] D. Gruhl, R. Guha, R. Kumar, J. Novak, and A. Tomkins, "The predictive power of online chatter," in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, pp. 78–87, San Diego, CA, USA, August 2005.

[39] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[40] W. Jabr and Z. Zheng, "Know yourself and know your enemy: an analysis of firm recommendations and consumer reviews in a competitive environment," *MIS Quarterly*, vol. 38, no. 3, pp. 635–654, 2014.

[41] X. Yu, Y. Liu, X. Huang, and A. An, "Mining online reviews for predicting sales performance: a case study in the movie domain," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 4, pp. 720–734, 2010.

[42] S. Ludwig, K. De Ruyter, M. Friedman, E. C. Brüggen, M. Wetzels, and G. Pfann, "More than words: the influence of affective content and linguistic style matches in online reviews on conversion rates," *Journal of Marketing*, vol. 77, no. 1, pp. 87–103, 2013.

[43] L. Zhang, *Python Data Analysis and Mining Practice*, Mechanical Industry Press, Beijing, China, 2016.

[44] D.-H. Park, J. Lee, and I. Han, "The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement," *International Journal of Electronic Commerce*, vol. 11, no. 4, pp. 125–148, 2007.

[45] S. Moon, Y. Park, and Y. S. Kim, "The impact of text product reviews on sales," *European Journal of Marketing*, vol. 48, pp. 2176–2197, 2014.

[46] J. A. Chevalier and D. Mayzlin, "The effect of word of mouth on sales: online book reviews," *Journal of Marketing Research*, vol. 43, no. 3, pp. 345–354, 2006.

[47] A. Floh, M. Koller, and A. Zauner, "Taking a deeper look at online reviews: the asymmetric effect of valence intensity on shopping behaviour," *Journal of Marketing Management*, vol. 29, no. 5-6, pp. 646–670, 2013.