

Research Article

Opinion Mining-Based Term Extraction Sentiment Classification Modeling

Tae Yeun Kim ¹ and Hyoung Ju Kim ²

¹National Program of Excellence in Software Center, Chosun University, Gwangju, Republic of Korea

²Department of Computer Engineering, Chosun University, Gwangju, Republic of Korea

Correspondence should be addressed to Hyoung Ju Kim; hyoungjukim@chosun.ac.kr

Received 24 January 2022; Accepted 2 April 2022; Published 27 April 2022

Academic Editor: Chun-Wei Tsai

Copyright © 2022 Tae Yeun Kim and Hyoung Ju Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The spread of social media has accelerated the formation and dissemination of user review data, which contain subjective opinions of users on products, in an e-commerce environment. Because these reviews significantly influence other users, opinion mining has garnered substantial attention in analyzing the positive and negative opinions of users and deriving solutions based on these analytical results. Terms that include sentimental information and used in user reviews serve as the most crucial element in sentimental classification. In this regard, it is crucial to distinguish the most influential terms in user reviews. This study proposed a document-level sentiment classification model based on the collection and application of user reviews generated in an e-commerce environment. Here, a term information extraction method was applied to the proposed model to select core terms, classify the selected terms according to parts of speech (POS), determine terms that can increase information power and influence, and adopt these terms in opinion mining research, based on SVM, SVM+, and SVM+MTL techniques. The results obtained from evaluating the proposed model indicate that it exhibited excellent sentiment analysis performance. The proposed model is expected to be effectively utilized in providing enhanced services for users and increasing competitiveness in the e-commerce environment.

1. Introduction

The development of information and communications technology has accelerated the advancement of social media, which facilitates communications among users, accumulates an enormous amount of knowledge, and enables users to easily exchange knowledge in online environments. Social media have provided online users with the ability to obtain and analyze data beyond the limitations of time and space. Users who purchase products in e-commerce environments often write reviews on these products, which serve as crucial sources of information for other online users who might be interested in purchasing products related to these reviews. In other words, these reviews are utilized as essential information for providing enhanced services to e-commerce users. However, owing to the rapid increase in the number of online reviews due to big data, it has become difficult to obtain useful information by analyzing these reviews. Hence, opinion mining based on text mining has garnered signifi-

cant attention, as this technique can present the results obtained from analyzing sentiments included in a considerable amount of user reviews [1–3].

Previous studies on opinion mining applied to online reviews primarily focus on classifying the reviews or opinions of users on products generated in an e-commerce environment, according to sentimental polarity (positive or negative sentiments) [4–6]. Certain online reviews might provide users with incorrect information, such as advertisements or intentionally distorted information. Therefore, research on opinion mining should utilize core terms that represent the subjective opinions of users in the reviews. In other words, it is necessary to extract core terms, which function as the most significant elements in sentiment classification, from the entire terms included in reviews and determine core terms that can represent documents. Hence, numerous studies have been conducted on opinion mining based on term information extraction [7–9]. Because the roles of terms including sentimental information perform differently according to POS in

these studies, such roles should be analyzed according to POS [10]. Nouns tend to reflect properties and sentiments related to products, while adjectives and verbs tend to include information on the subjective opinions and evaluation of users. In addition, adverbs are applied as various types of expression methods and modification functions in the process of document classification [11, 12].

Existing studies have also adopted data mining techniques that can predict the sentiments of users by using text documents to analyze their sentiments based on opinion mining from a different perspective [13–15]. Although several researchers have utilized data mining techniques to derive enhanced prediction performance, they have failed in addressing the inherent limitations of these techniques. To address this problem, a number of studies improved prediction performance by integrating data mining techniques. Increasing attention has also been paid to a research, which proposed an integrated model considering the characteristics of each data mining technique to achieve outstanding prediction performance instead of performing the duplicated application of a simple model.

The objectives of this study are outlined as follows: First, in sentence-level analysis, a number of studies have been carried out that utilized various parts of speech to determine the main part of speech (POS) and conduct sentiment analysis. In particular, based on the approach of semantic analysis, adjectives bear the most sentiment information among different POS, and they are used as the core terms for classification of polarity. However, in a document-level analysis, optimal selection of terms is instrumental since selection of input variables determines the predictive performance of the model and has a significant impact on the classification of the sentiment of the entire document. Also, it has been reported that considering the entire document using various POS is more effective than considering only adjectives in a fragmented approach. Therefore, in this study, terms are extracted focusing on the four POS (adjectives, adverbs, verbs, and nouns) and the extracted terms are used as input variables for a model for document-level sentiment classification. Also, considering the superiority of adjectives in terms of sentiment bearing, we aim to examine whether there are differences in the relationship between the four POS and the prediction performance of different model types.

Second, in the document-level analysis, using an optimal number of terms as input variables for the model is important for effective classification of the document. If too many terms are used as input variables, it is highly likely that there will be missing values and high redundancy; thus, it is difficult to determine if sentiment information is included, and also, the efficiency of the analysis is reduced. Therefore, in this study, a variety of term information extraction techniques are used in order to select effective and efficient terms as input variables to be used in the model and compare the results and performance between different techniques.

Third, in opinion mining research, when data mining techniques are used, the selection of the variables used in the model and the method applied according to the characteristics of the data plays the central role. In addition, since different results are obtained depending on the use of vari-

ous data mining techniques, the model performance may vary. That is, it is important to explore and select the most efficient model by adopting various data mining techniques. Therefore, in this study, SVM+ and SVM+MTL techniques, which have not been implemented in opinion mining research to date, are applied, and models for document-level sentiment classification are developed to compare the prediction performance of the developed models.

Lastly, in opinion mining research, different single models are combined to develop a new integrated model, which serves as a method for further improvement of the predictive performance of a single model. In general, an integrated model is applied with various methods to improve the performance of a single model, and a number of studies have been carried out on the development and application of the integrated models. Thus, in this study, through the results of the models for document-level sentiment classification using SVM, SVM+, and SVM+MTL techniques, we propose an integrated data mining model, a type of an integrated model, and comparatively analyze the difference in predictive performance between the single model and the integrated model.

To this end, core terms were selected by adopting a term information extraction technique, and then, the selected terms were classified according to POS to determine the final terms that can increase information power and influence. Subsequently, approach presented an integrated sentiment classification model by applying SVM+ and SVM+MTL techniques to the research on opinion mining. To evaluate the performance of the proposed model, this study collected 80 000 user reviews on movies, games, music, and books on Amazon (<http://Amazon.com>), eliminated unnecessary terms, and extracted terms based on POS tagging. It also calculated the values of document frequency, term frequency-inverse document frequency (TF-IDF), information gain, and chi-square statistic to rank the extracted terms based on the values derived. Subsequently, the top 20 optimal terms were determined according to the categories and classified sentiments included in the optimal terms selected by utilizing SVM, SVM+, and SVM+MTL techniques. The obtained evaluation results indicate that the sentiment classification model based on the chi-square statistic exhibited the most excellent prediction performance among the other models. In addition, the proposed sentimental classification model effectively analyzed sentiments included in terms. It is expected that the proposed term-based sentiment classification model can be effectively used to improve services, ensure competitiveness, and provide enhanced services for users in the e-commerce environment. The remainder of this paper is presented as follows: Section 2 reviews previous studies, while Section 3 introduces the proposed sentiment classification model based on terms at the document level. Furthermore, Section 4 presents the results obtained from conducting experiments on the proposed model. Finally, Section 5 provides conclusions and future research directions.

2. Related Research

2.1. Opinion Mining for Sentiment Classification. Opinion mining, which was developed from a data mining technique for document classification, refers to the process of

extracting, classifying, and analyzing the opinions of users expressed in various media [16]. This technique, also called sentiment analysis, analyzes sentiments expressed in text and a state of positive or negative expressions on a certain subject included in the text. Owing to such characteristics, the field of opinion mining for social network analysis has garnered increasing interests [17, 18]. Sentiment analysis generally determines the positive or negative tendency of an entire document by comprehensively analyzing opinions included in the document. To achieve this objective, the following analyses are performed at two different levels [19]: First, a document-level analysis is conducted to classify documents that express positive or negative opinions or sentiments. In this analysis, opinions on a certain subject are regarded as the basic unit of information. This analysis facilitates fast and comprehensive determination, as certain terms included in a document are selected according to frequency. Second, a sentence-level analysis is conducted to classify sentiments expressed in sentences that are included in a document. This analysis enables people to determine whether a positive or negative opinion is expressed, based on sentences in the document. Specifically, this analysis technique determines an opinion based on sentences, extracts sentences including sentiments, and examines paragraphs or phrases, considering different terms [20].

Sentiment classification is the process of extracting expressional terms in a document or a sentence from reviews or comments, identifying sentiments in various forms of documents, and classifying these sentiments according to polarity (positive or negative sentiments) [21]. Studies on sentiment classification adopt an approach based on vocabulary, machine learning (ML), or a method integrating vocabulary and ML [19]. In the vocabulary-based approach, a model analyzes the intentions of users based on positive and negative terms and the implications of these terms. In this process, it primarily adopts previously edited or well-known sentiment terms [22, 23]. In the ML-based approach, a model trains behavior or pattern data based on data collected and then predicts the future by classifying or analyzing these data. This process applies algorithms and linguistic variables used in the ML-based approach and depends on algorithms to solve document classification problems related to sentiment analysis. In the integrated approach, a model adopts both the vocabulary- and ML-based approaches. In general, sentiment terms serve as crucial elements in this approach. Figure 1 illustrates various approach techniques for sentiment classification.

2.2. Term Information Extraction Technique. Term information extraction is the most significant methodology for determining terms that contain the greatest amount of information on document classification and establishing a classification model to solve sentiment classification problems at the document level in opinion mining research [24, 25]. Terms used in reviews or comments contain sufficient information that can be adopted to classify linguistic properties and documents, thereby performing a crucial role in sentiment classification, and serving as an important basis for determining the main sentiment in a review. It is crucial to

determine core terms, as these terms exert significant influence on determining the polarity of the entire document.

2.2.1. Document Frequency. Document frequency refers to the number of documents that include certain terms among the entire documents. In other words, it is a ratio of documents that contain terms that were used over a certain frequency in documents. This technique is simple and requires less calculation than other techniques. It also improves the accuracy of document categorization by eliminating low-frequency terms and discarding words under the assumption that terms that occur at low frequency do not contribute to document classification [26, 27]. Equation (1) defines document frequency, where N and i denote the entire number of documents and terms, respectively.

$$DF_i = \frac{n_i}{N}. \quad (1)$$

2.2.2. Information Gain. Information gain is one of the most frequently used techniques in opinion mining research, which is known to be more influential than document frequency. This technique calculates the amount of information included in terms according to categories by considering both the occurrence and absence frequencies of terms, respectively. It is applied as an evaluation standard for measuring the usefulness of terms in the ML field. Information gain is calculated via the following process. First, a set of the entire documents is divided into subsets. Second, entropy, which changes before and after the use of a certain term, is calculated. Third, the probability of a target is divided and calculated according to categories, to derive an information value on the certain terms. Fourth, the gain is calculated to derive normalized gain [28, 29]. Equation (2) defines information gain, where t denotes a term.

$$IG(t) = \sum_{i=1}^c P(c_i) \log \frac{1}{P(c_i)} - \sum_{t \in \{t_p, \bar{t}_p\}} P(t) \sum_{i=1}^c P(t | c_i) \log \frac{1}{P(t | c_i)}, \quad (2)$$

where t_p is the number of times the term t occurs in the document and \bar{t}_p is the number of times the term t does not occur in the document.

2.2.3. TF-IDF. TF-IDF is a method that can represent the significance of a term in individual documents. This method is frequently adopted because of its advantages, which includes its similarity with document frequency, simple calculation processes, and excellent performance [30]. The significance of a term is proportional to the number of times the term appears in a document and inversely proportional to the number of entire documents that contain the term [31, 32]. The TF-IDF technique was developed as a document representation method to evaluate the relative significance of terms in a document. This technique is frequently used in analyses based on the unit of a phrase or a paragraph [33]. It calculates the significance of a term by applying a

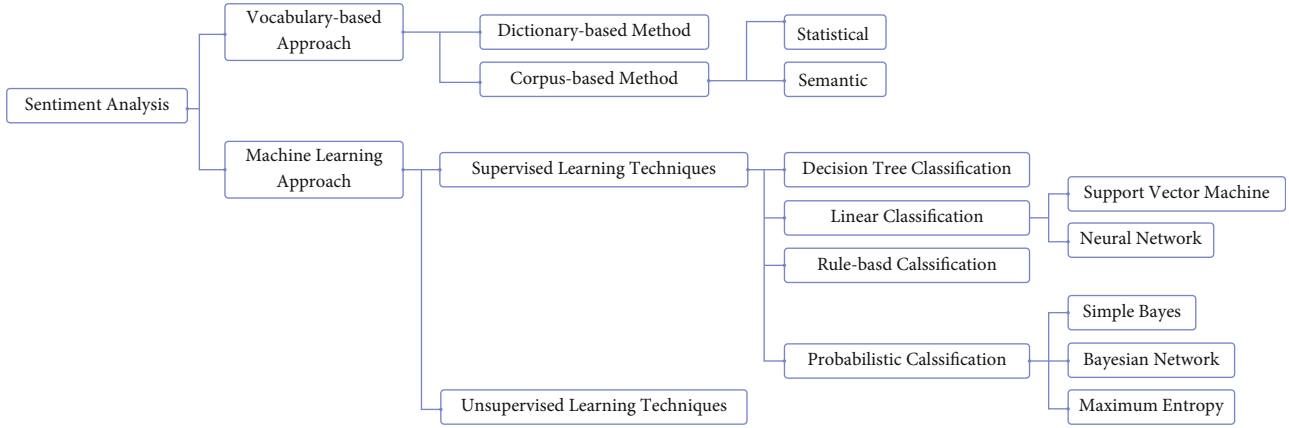


FIGURE 1: Sentiment classification technique.

weight to the term and employing TF, which refers to the number of times a word appears in a document, and DF, which is the number of the entire documents containing the term. Equations (3) and (4) express the TF-IDF calculation processes based on a term (t) and document (d).

$$TF - IDF_{t,d} = TF_{t,d} * IDF_t, \quad (3)$$

where

$$IDF_t = \log \frac{N}{DF_t}, \quad (4)$$

where N denotes the total number of documents, $TF_{t,d}$ the number of documents (d) and terms (t) that appear, and DF_t the total number of documents containing the term (t).

TF alone cannot be used to calculate the significance of a term because a high-frequency term is likely to be meaningless in a document. Consequently, both TF and IDF are reflected in the calculation process. IDF divides the number of the entire documents by the number of documents that contain the given term. Therefore, a term that appears in numerous documents has a low IDF value. In contrast, a term that is biased to certain documents has a high IDF value.

2.2.4. Chi-Square Statistic. The chi-square statistic, which analyzes a correlation between categorical variables, is the most frequently used technique among cross-tabulation analysis techniques. Specifically, the value of a correlation is “1” if a term appears more than once in the entire documents; otherwise, it is “0.” In other words, this technique is used to identify the significance between a term (t_i) and a category (C_j).

$$x^2 = \frac{N \times (AD - CB)^2}{(A + C) \times (B + D) \times (A + B) \times (C + D)}, \quad (5)$$

where

$N = A + B + C + D$ (N the total number of documents),
 $i = 1, 2, \dots, n$ (n the total number of terms used),
 $j = 1, 2, \dots, m$ (m the total number of categories).

The chi-square statistic is similar to mutual information; however, the former exhibits better performance for term extraction than the latter, owing to the standardized value of the former. Equation (5) defines the chi-square statistic (x^2) between C_j and t_i in a range of the consistent properties applied. Table 1 presents the document frequency between C_j and t_i [34, 35].

2.3. Data Mining Techniques

2.3.1. SVM. The SVM technique, which was developed by Vapnik based on the statistical learning theory, exhibits outstanding performance in solving classification problems [36]. While most learning algorithms focus on empirical risk minimization, this technique aims at structural risk minimization. Because this technique can easily generalize classification problems, it has been applied in various fields. The SVM technique can also match nonlinear problems on input space with linear problems on specific high-dimensional space, thereby facilitating convenient mathematical analysis [37, 38]. Kernel functions generally used for the SVM technique include the polynomial kernel and Gaussian radial basis function (RBF). These functions form input datasets that can be linearly separated in specific space by mapping data in high-dimensional space [39]. Figure 2 illustrates the decision boundary and margin of the SVM technique.

2.3.2. SVM+. Vapnik also proposed SVM+, which is an optimal SVM-based technique for managing learning with structured data (LWSD) or learning using privileged information (LUPI) to obtain hidden group data from training data [36]. In addition, Vapnik defined a slack variable in each group based on a modification function by considering group data and then presented a method of mapping input vectors in two different Hilbert spaces simultaneously. Figure 3 presents the mapping method.

The slack variable is restricted by a correction variable under the applied SVM+ environment. A mapping sample in a modification space should be placed on one side of a correspondence function. The SVM+ technique simultaneously maps data in both decision and modification spaces. When a

TABLE 1: Document frequency between term (t_i) and category (C_j).

	Documents in which category (C_j) occurred	Documents where category (C_j) does not occur
Documents with the term (t_i)	A	B
Documents that do not have the term (t_i)	C	D

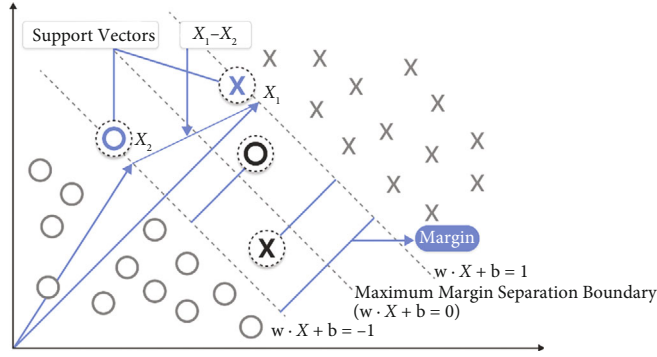


FIGURE 2: Decision boundary and margin of the SVM.

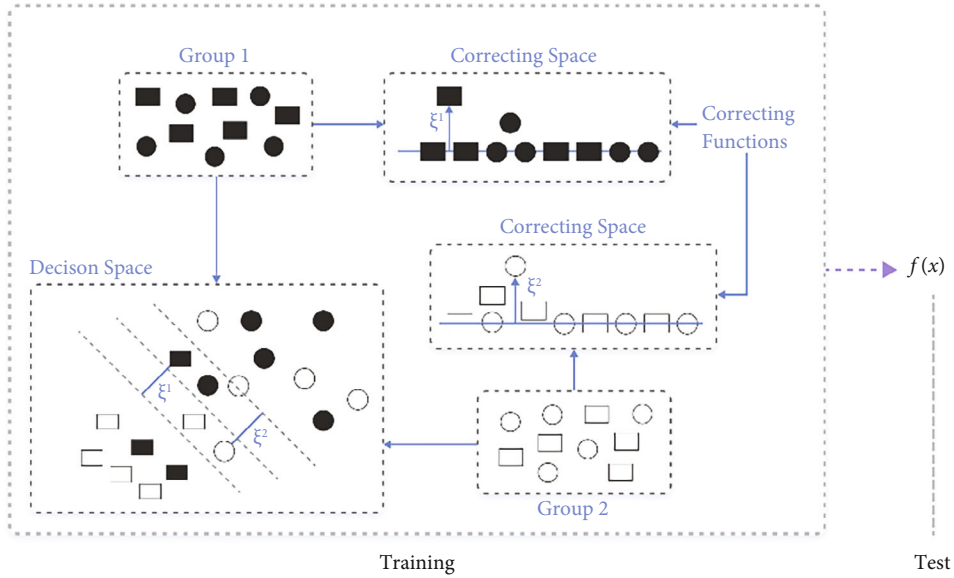


FIGURE 3: Processing using the SVM+.

slack modification function is defined in the same decision space, a decision function is defined in the decision space. Data in other groups are mapped in the same decision space, whereas a modification function can be defined in the same modification space or a different modification space, owing to the application of different correction variables in different groups [40]. Equations (6) and (7) define the SVM+ technique:

$$\min_{W, b \in R, \xi^r} \phi(W, b, \xi^r) = \frac{\|W\|^2}{2} + \frac{r}{2} \|W_r\|^2 + C \sum_{r=1}^t \sum_{i \in T_r} \xi_i^r \quad (6)$$

Subject to:

$$y_i (W^T \phi(X_i + b)) \geq 1 - \xi^r, i \in T_r, r = 1, \dots, t \quad (7)$$

$$\xi^r \geq 0, i \in T_r, r = 1, \dots, t \quad (8)$$

$$\xi^r = (W^T \phi_i(X_i) + d_r), i \in T_r, r = 1, \dots, t. \quad (9)$$

R denotes the number of groups, while y adjusts the capacity of modification space and importance of capacities of the modification space. ξ^r represents the slack variables of each group and differs according to groups. A penalty coefficient (C) maintains a balance between the complexity of a model

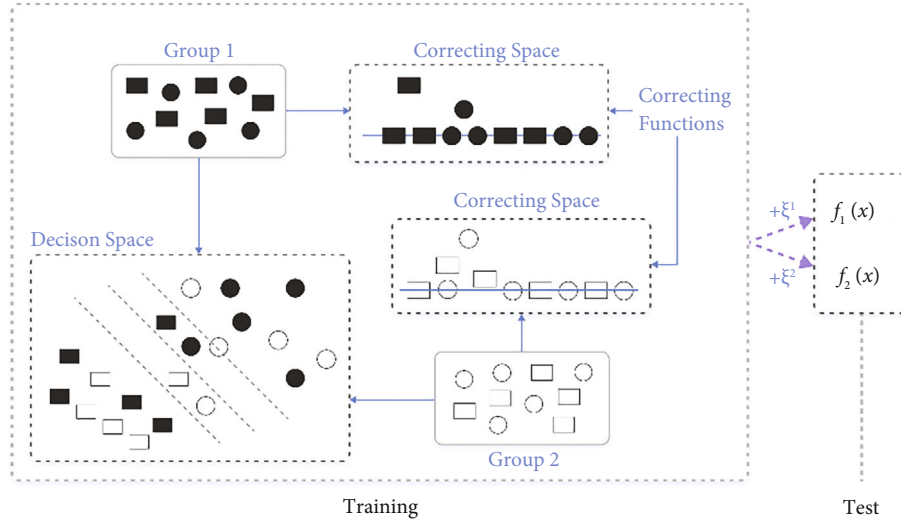


FIGURE 4: Processing using the SVM+MTL.

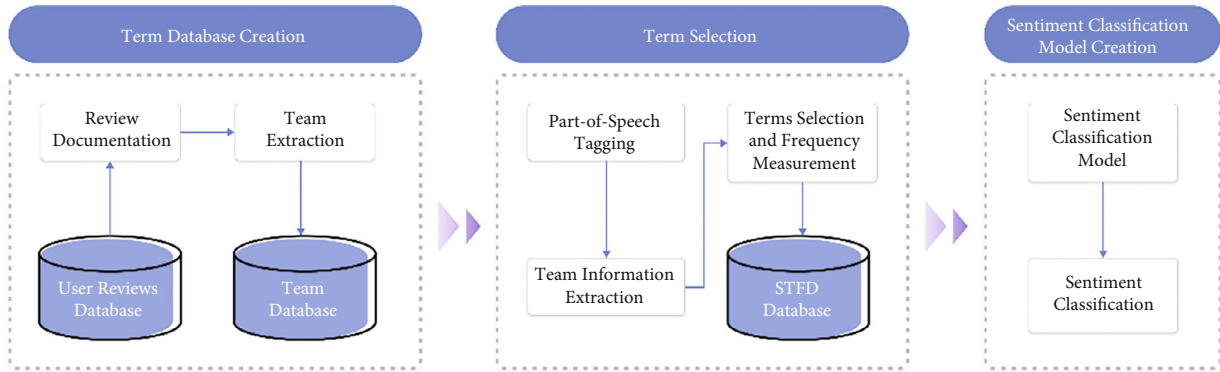


FIGURE 5: Term-based sentiment classification model at the document level.

(i.e., training capability) and the error rate of a sample. That is, the SVM+ technique can be divided into two parts, where W denotes the capability of the decision space and W_r denotes the capability of the crystal space. However, since W_r does not determine the size of the margin, the SVM+ technique adds $(r/2)\|W_r\|^2$ compared to the SVM technique.

The specific term information extraction processes of the SVM+ technique are consistent with those of the SVM technique. In the term information extraction process, special attention should be paid to the following issues:

- (1) The SVM+ technique should divide training data by selecting appropriate group data according to professional or general knowledge, instead of performing statistical analysis
- (2) The amount of samples can significantly influence the response effects of the SVM+ technique. Hence, the SVM technique might exhibit the best perfor-

mance when the amount of samples is small, owing to a small number of parameters to be adjusted

- (3) The SVM+ technique can significantly influence the selection of both kernel functions and parameters
- (4) The SVM+ technique is consistent with the SVM technique when each training vector among the entire data belongs to independent groups that are not related to each other
- (5) The SVM+ technique requires a longer time for experimental calculation than the SVM technique

Ribeiro et al. adopted the SVM+ method as a model for predicting corporate bankruptcy based on financial data obtained in France and then examined the prediction performance of the model by considering authority information [41]. Specifically, they analyzed the prediction performance of the corporate bankruptcy prediction model by using

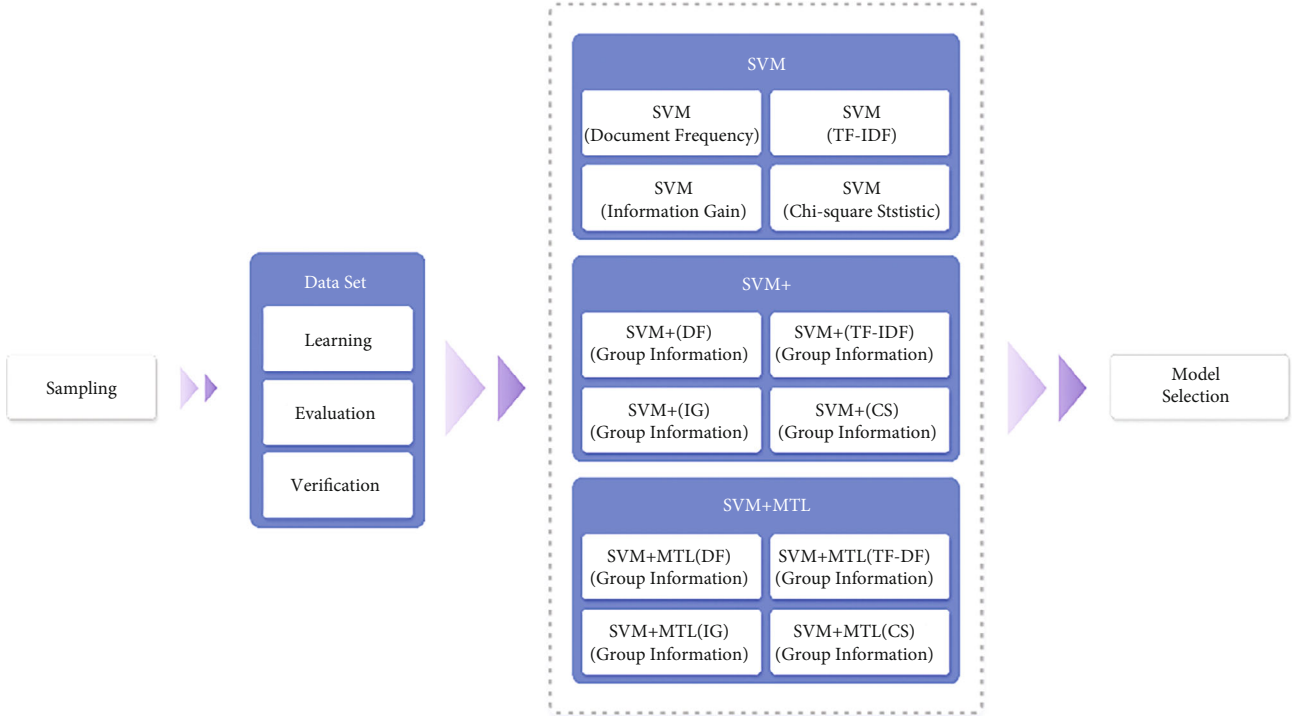


FIGURE 6: Sentiment classification model.

heterogeneous information grouped according to the size categories of 30 companies relative to financial ratios, number of employees, and annual profits. Based on the results obtained from their analyses, they argued that the SVM+ technique always exhibits a more optimal performance than the SVM and SVM+MTL techniques, regardless of kernel functions. Moreover, the corporate bankruptcy prediction model that applies the Gaussian RBF as a kernel function exhibited better performance than the model that applies a linear kernel function based on the application of structured information in training data.

Serra-Toro et al. elucidated the relationship between LUPI and the SVM+ technique, thereby indicating that the performance of this technique depends on a subtle relationship between normal data and authority information [42]. They also stated that a randomly generated variable can also perform a crucial role when it contains authority information on a specific problem.

In addition, they argued that the performance of the SVM+ technique can exceed that of the SVM technique according to data preprocessing, data set segmentation, validation protocols, experiments, and a range of parameters and search procedures.

2.3.3. SVM+MTL. Liang et al. developed an SVM+MTL technique by combining the SVM+ algorithm and the multitask learning (MLT) technique. Specifically, the SVM+ technique was adopted to define different decision functions for each group and identify a relationship between groups [43]. To solve problems related to the MTL technique, the

SVM+MTL technique also maps data in two Hilbert spaces simultaneously, similar to the SVM+ technique. Equation (10) defines the decision function.

$$f_r(x) = W^T \phi_Z(X_i) + b + W_r^T \phi_{Z_r}(X_i) + d_r, r = 1, \dots, t. \quad (10)$$

The SVM+MTL technique has the following advantages over the SVM+ technique:

- (1) Group data also exist in prediction data
- (2) A slack variable indicates an error in the entire model after a modification function is added, instead of being defined by the modification function
- (3) A decision function has a correction term unlike those used for the SVM and SVM+ techniques

Liang et al. proposed an MTL algorithm for the SVM+MTL technique. This study utilized group data to solve MTL-related problems by applying the SVM+ technique [43]. Via this process, the prediction accuracy of the proposed model increased. Moreover, this study elucidated the relationships between SVM+ and SVM+MTL techniques and compared the performance of sentiment classification models that apply these techniques by using comprehensive datasets to empirically compare the performance of these models. The results obtained from the comparison indicated that the sentiment classification model that applies the SVM+MTL technique exhibited excellent prediction performance when the amount of data was sufficient. However, when the

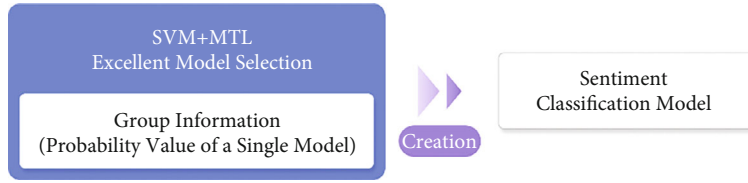


FIGURE 7: Sentiment classification model construction.

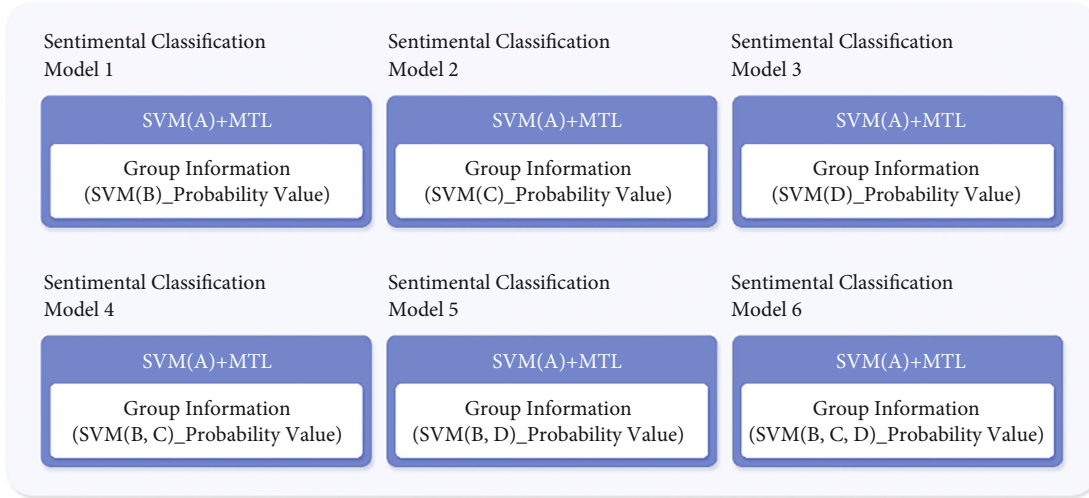


FIGURE 8: Application result of sentiment classification model.

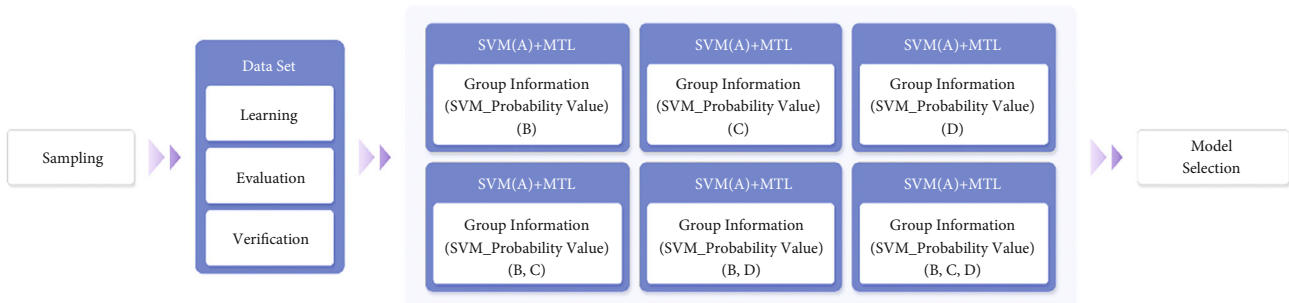


FIGURE 9: Document-level sentiment classification model.

amount of data was insufficient, sentiment classification models that apply the SVM or SVM+ technique exhibited outstanding performance. Figure 4 presents processing using the SVM+MTL..

3. Sentiment Classification Modeling

3.1. *Sentiment Classification Model.* The sentiment classification model proposed in this study determines and classifies core terms, which contain a significant amount of information and influence the entire document, for prediction. Figure 5 illustrates the three stages required to establish a term-based sentiment classification model at the document level, which include the formation of a term database (DB),

selection of terms, and establishment of a sentiment classification model.

3.1.1. *Processes of a Term DB Formation.* The processes of a term DB formation are as follows: First, customer review data generated on the web are crawled, collected, and stored in a customer review DB.

Second, these data are preprocessed, owing to the unstructured data. In general, unstructured data obtained from the web should be converted to structured data to be used in an experiment. In other words, a customer review can accurately reflect a subjective opinion of an individual only when this review is considered as a document.

Third, structured data obtained via preprocessing are adopted to classify reviews as positive and negative reviews.

TABLE 2: Number of terms extracted by item.

	Movie	Game	Music	Book
Number of documents	30,000	30,000	20,000	20,000
Number of terms collected	77,222	78,523	66,190	85,456
Adjective	132	146	117	140
Noun	400	402	319	309
Adverb	57	55	61	63
Verb	161	147	169	154

This classification process is required to identify positive or negative tendencies of customers in their opinions. Hence, this study developed a simple application program based on the Microsoft foundation class (MFC) to document individual customer reviews.

Fourth, the entire terms included in the documents are extracted. The extracted terms contain stop words, which appear frequently, exhibit weak information power, and do not provide special information. Hence, it is necessary to eliminate such stop words. Stop words that appear frequently in a document consist of articles (e.g., a, an, and the), conjunctions (e.g., that, and, and when), pronouns (e.g., I and you), and the verb to be (e.g., is and are). The process of eliminating stop words is performed to select highly influential terms, develop a term DB based on the selected terms, and store the DB.

3.1.2. Term Selection. Subsequently, four POS (adjectives, adverbs, verbs, and nouns) are utilized to tag the terms stored in the term DB. Information on POS serves as a crucial index for deriving sentimental information from terms and can also be effectively used to determine terms with significant information power and influence by accurately analyzing the characteristics of terms. However, when POS tagging is not performed, the term-selection process requires numerous manual tasks, including term selection based on the opinions of experts and a considerable amount of cost and time.

In this study, an experiment was conducted in which a Stanford POS Tagger program, a grammatical structure analysis tool, was used to apply 4 POS tagging to the entire terms and measure the frequency of occurrence of these terms in the entire documents. Subsequently, term information extraction techniques were adopted to calculate term information values. Document frequency, TF-IDF, information gain, and chi-square statistic were applied as term information extraction techniques in this study. After the information values of each term were calculated, these terms were ranked according to term information extraction techniques to determine the terms to be used as input variables and store the selected terms in a sentiment terms for documents database (STFD).

3.1.3. Sentiment Classification Model Formation. The selected terms presented in Figure 6 are adopted as input variables, and the dataset including these terms is divided

into three different datasets for training, evaluation, and verification. Next, SVM, SVM+, and SVM+MTL techniques for term information extraction are applied to establish document-level sentiment classification models.

3.2. Term-Based Sentiment Classification Model. The term-based sentiment classification model proposed in this study is an integrated model that exhibits better performance than individual sentiment classification models, beyond the limitations of these models.

Figure 6 presents the integrated sentiment classification model based on the SVM+MTL model, which reflects the experimental results of individual sentiment classification models based on the aforementioned term information extraction techniques, to exhibit more excellent sentiment analysis performance. The integrated model based on the SVM+MTL technique utilizes and trains group data as probability values (prediction performance rates) obtained in the experimental results on individual sentiment classification models.

Figure 7 presents the processes of establishing the integrated sentiment classification model proposed in this study. These processes are as follows:

- (1) The experimental results based on sentiment classification models applying SVM, SVM+, and SVM+MTL techniques are derived, as illustrated in Figure 6
- (2) The model, which exhibited the best result among sentiment classification models that apply the SVM+MTL technique, is selected
- (3) Group data used for the SVM+MTL technique are used as probability values for the remaining three models, except for the best model
- (4) Models are divided into specific models with group data applied in various ways for training, as illustrated in Figure 8
- (5) Prediction results derived by the sentiment classification model are applied to the decision-making processes

When A, B, C, and D refer to document frequency, TF-IDF, information gain, and the chi-square statistic, respectively, and a model based on A exhibits the best performance in the experimental results based on sentiment classification models that apply the SVM+MTL technique, the sentiment classification model applying the SVM+MTL technique based on A is selected to establish the integrated sentiment classification model. Accordingly, probability values of sentiment classification models that apply the SVM technique based on B, C, and D are calculated and adopted as group data. Consequently, sentiment classification models are subdivided and defined. Figure 8 illustrates the defined models.

Figure 9 presents document-level sentiment classification models. Accordingly, sentiment classification models 1, 2, 3, 4, 5, and 6 are generated.

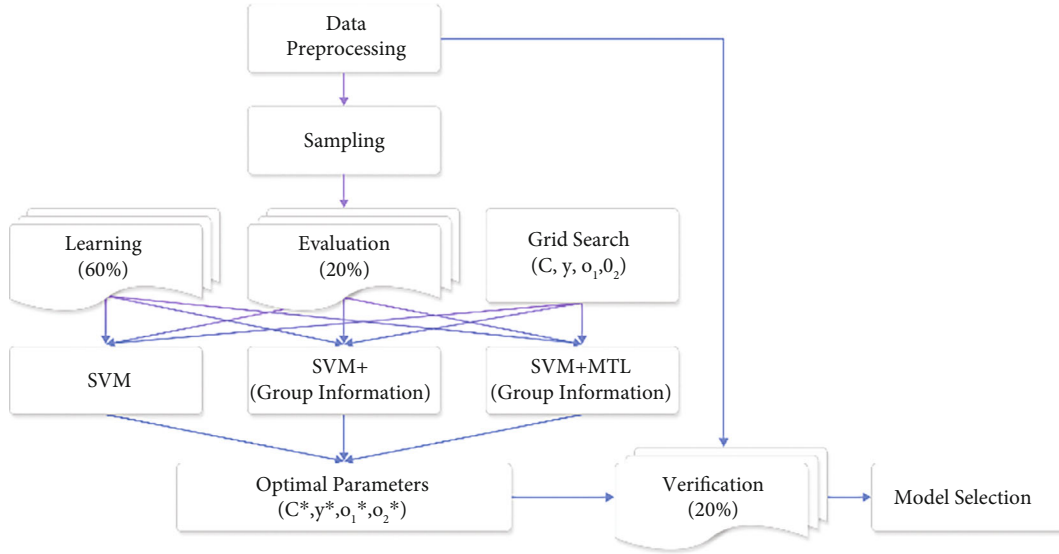


FIGURE 10: Single model experiment process for sentiment classification.

TABLE 3: Confusion matrix.

	Predicted positive	Predicted negative
Actual positive	TP (true positive)	FN (false negative)
Actual negative	FP (false positive)	TN (true negative)

4. Sentiment Classification Model Experiment and Results

4.1. Experiment Data. This study employed user review data generated on Amazon (<http://amazon.com>). Particularly, it collected user review data associated with the categories of movies, games, music, and books, in which users performed enormous purchase activities for relevant products. The data of these user reviews were directly crawled from July 2020 to August 2021, and MFC was utilized to document each user review. Accordingly, an application program for document classification was developed and utilized. First, the list of each item was formed. Then, documents corresponding to Score 4 or higher and documents corresponding to Score 2 or lower were classified and collected as positive and negative documents, respectively. Because documents corresponding to Score 3 were classified as neutral documents, they were excluded. Specifically, 150 movie reviews, 150 games reviews, 126 music reviews, and 134 book reviews were collected. For movie and game reviews, 30 000 review documents were generated for each category. Regarding music and book reviews, 20 000 review documents were generated for each category. A program R was utilized to calculate the occurrence frequency of the remaining terms, except stop words in the entire documents.

Table 2 presents the number of documents for each category, number of terms collected via POS tagging performed by the Stanford POS Tagger program, and results obtained from extracting 1000 terms of 4 POS that appeared frequently. Term information extraction techniques were

applied to calculate term information values, and the terms were ranked according to each technique. Via this process, this study selected 20 terms of 4 POS that appeared frequently, according to each technique.

Numerous input variables can trigger problems such as an increase in missing values and difficulties in model management. Therefore, a model with a low number of parameters for input variables is more affected by input variables. A significant number of input variables can also generate variables that are correlated to each other. The elimination of these variables reduces the variance of prediction values, thereby enhancing a model with increased prediction power. Hence, this study selected 20 input variables, which had significant effects on classifying the properties of the entire documents and contained the greatest amount of information on document classification, to be used in the document-level analysis.

The experimental environment of this study is described as follows. Experiments were carried out in “Matlab” to calculate and evaluate experimental results. The specifications of the computer used in this experiment are as follows: Windows 10, Intel® Core™ i5-10500 CPU@3.10GHz 3.10GHz processor, 16.00-GB RAM, and a 64-bit operating system.

4.2. Experiment Design

4.2.1. Term-Based Sentiment Classification Model Experiment Design. To apply balanced data in this experiment, the entire datasets were classified as positive and negative documents. Document-level sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques adopt the Gaussian RBF kernel function and grid search to identify optimal parameters. C denotes the upper limit value and γ represents a parameter that adjusts the weight between decision and modification function, while σ_1 and σ_2 are parameters the Gaussian RBF kernel function in the decision and modification functions, respectively. The

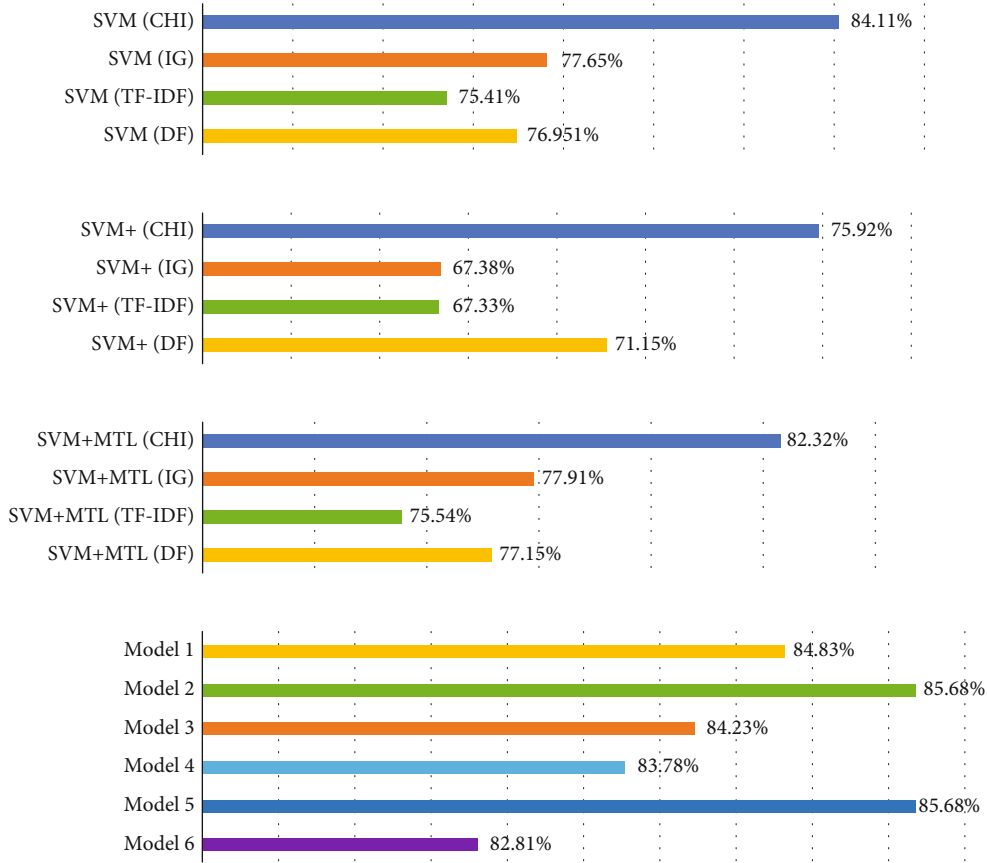


FIGURE 11: Comparison of overall model performance in movie reviews.

parameters used in this experiment were established as follows: $C = \{1, 25, 50, 75, 100\}$, $\gamma = \{0.001, 0.01, 0.1, 1, 10\}$, and $\sigma_1, \sigma_2 = \{1, 5, 7, 8, 10\}$.

Regarding the group data adopted by the sentiment classification model that applies the SVM+ technique and the sentiment classification model that applies the SVM+MTL technique, information on the year of issue was employed according to items. The confusion matrix was applied to calculate prediction results derived by document-level sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques. Figure 10 presents single model experiment process for sentiment classification.

As presented in Table 3, the confusion matrix shows a relationship between the actual categories and those predicted and classified by models. Accuracy refers to the accuracy of sentiment classification results. With an increase in accuracy, the model improves in its performance. Reproducibility is a ratio of the actual values to the existing sentiment classification results, while precision refers to the accuracy of the prediction. Equations (11), (12), and (13) define accuracy, reproducibility, and precision, respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (12)$$

$$\text{Precision} = \frac{TP}{TP + FP}. \quad (13)$$

F-measure is a method that evaluates the validity of actual and predicted data by utilizing the confusion matrix to evaluate the performance of a model in the field of data mining. This method can measure classification performance by considering both precision and reproducibility, including reflecting effects triggered when actual data are inclined to a certain target. It is frequently utilized to verify sentiment classification performance. Equation (14) expresses the process of calculating F-measure:

$$\text{F-measure} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}. \quad (14)$$

4.3. Experiment Results. The results obtained from analyzing each sentiment classification model based on movie reviews in experiments are presented as follows: the sentiment classification models that apply the SVM, SVM+, and SVM+MTL techniques based on the chi-square statistic exhibited prediction accuracies of 84.11%, 75.92%, and 82.32%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM+MTL technique exhibited the best prediction accuracy among the other single models. Sentiment classification models 2 and 5, which are integrated models proposed in

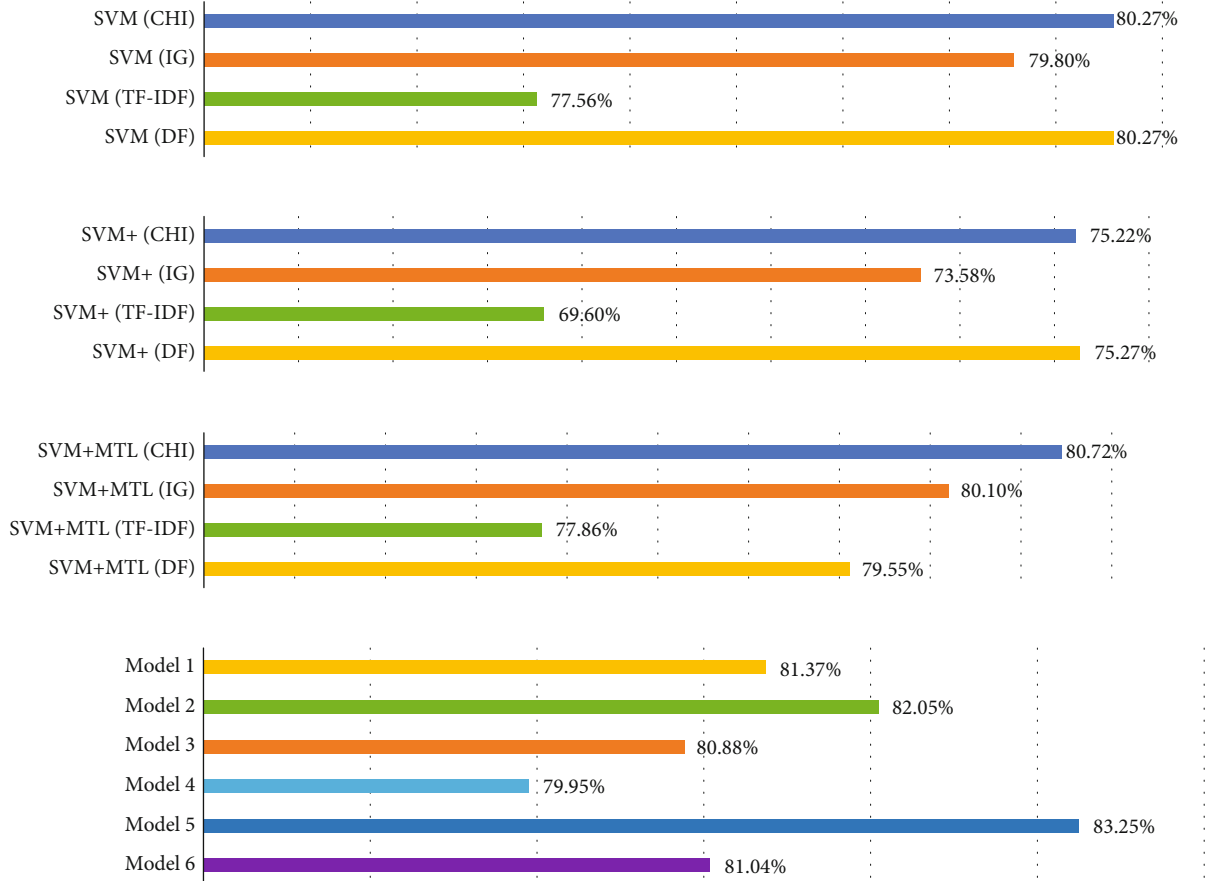


FIGURE 12: Comparison of overall model performance in game reviews.

this study, both exhibited a prediction accuracy of 85.68%, thereby exhibiting better prediction performances than single models. Figure 11 presents the results obtained from comparing the performance of the entire models based on movie reviews.

The results obtained from analyzing each sentiment classification model based on game reviews in the experiments are presented as follows: the sentiment classification models that apply the SVM technique based on the chi-square statistic and document frequency both exhibited a prediction accuracy of 80.27%, while the sentiment classification models that apply the SVM+ and SVM+MTL techniques based on the document frequency and chi-square statistic exhibited prediction accuracies of 75.27% and 80.72%, respectively. Based on these analytical results, it was verified that the sentiment classification model that apply the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 5, which is an integrated model proposed in this study, exhibited a prediction accuracy of 82.33%, thereby indicating that it exhibited a better prediction performance than single models. Figure 12 presents the results obtained from comparing the performances of the entire models based on game reviews.

The results obtained from analyzing each sentiment classification model based on music reviews in the experi-

ments are presented as follows. The sentiment classification models that apply the SVM, SVM+, and SVM+MTL techniques based on the chi-square statistic, TF-IDF, and chi-square statistic exhibited prediction accuracies of 80.45%, 75.23%, and 79.85%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 4, which is an integrated model proposed in this study, exhibited a prediction accuracy of 80.96%, thereby exhibiting a better prediction performance than single models. Figure 13 presents the results obtained from comparing the performances of the entire models based on music reviews.

The results obtained from analyzing each sentiment classification model based on book reviews in the experiments are presented as follows: the sentiment classification model that applies the SVM, SVM+, and SVM+MTL techniques based on the chi-square statistic, document frequency, and chi-square statistic exhibited prediction accuracies of 80.93%, 72.29%, and 81.00%, respectively. Based on these analytical results, it was verified that the sentiment classification model that applies the SVM+MTL technique based on the chi-square statistic exhibited the best prediction accuracy among the other single models. The sentiment classification model 6, which is an integrated model proposed in this

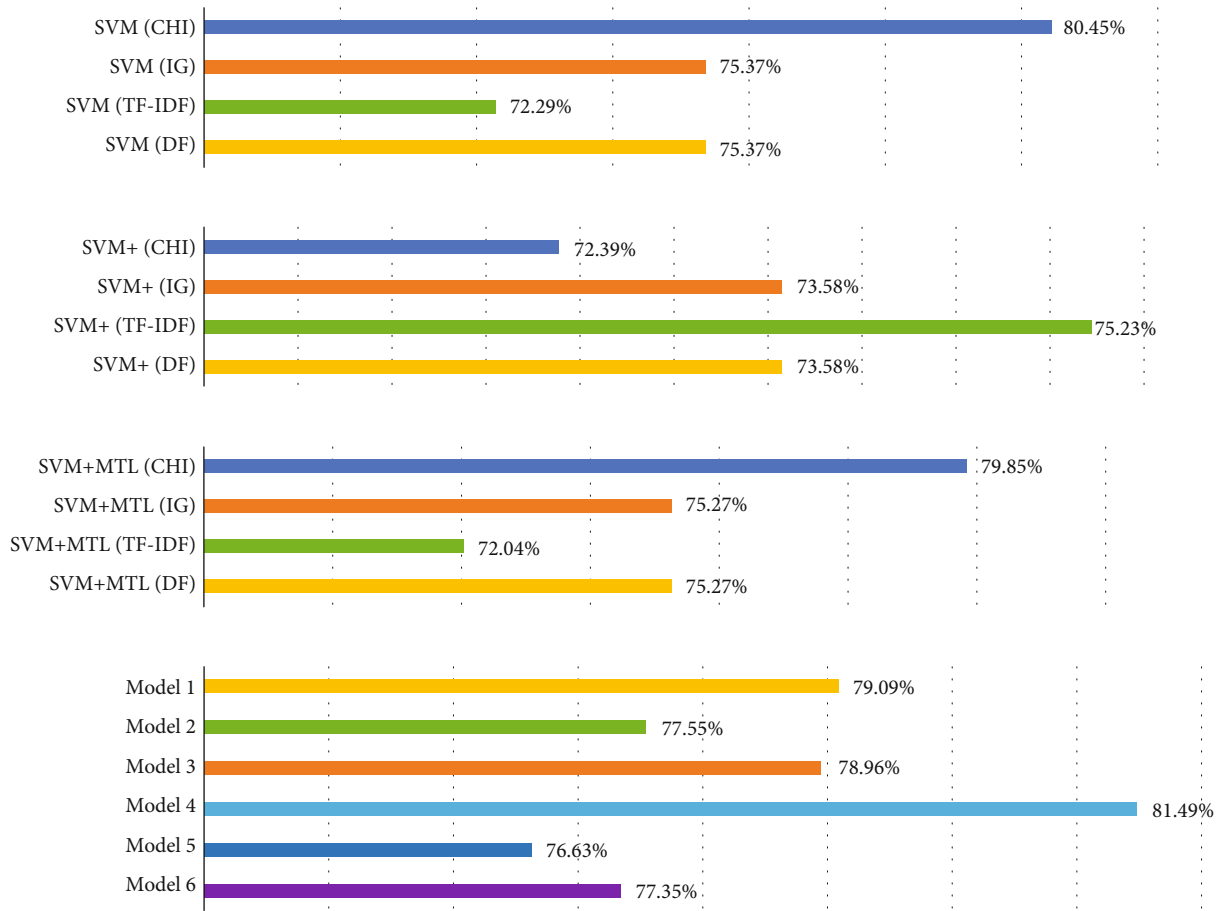


FIGURE 13: Comparison of overall model performance in music reviews.

study, exhibited a prediction accuracy of 81.11%, thereby exhibiting a more improved prediction performance than single models. Figure 14 presents the results obtained from comparing the performance of the entire models based on book reviews.

Tables 4 and 5 present the results obtained from analyzing sentiment classification models based on four cases (i.e., movies, games, music, and books). The analytical results indicate that the sentiment classification model that applies the chi-square statistic exhibited the most effective performance in terms of extracting core terms. This result was obtained because the chi-square statistic exhibited a close relationship with documents and terms. In other words, because this technique is closely related to terms with significant information power, the significance of these terms was reflected in documents.

This study predicted the performance of sentiment classification models that apply SVM, SVM+, and SVM+MTL techniques according to the amount of data and application of techniques. It was inferred that the sentiment classification model proposed in this study exhibited the best performance among the others. In addition, because the sentiment classification model that applies the SVM+MTL technique adopted group data for training, verification, and evaluation,

it exhibited better performance than the sentiment classification models that apply the SVM and SVM+ techniques, respectively.

It was expected that the sentiment classification model that applies the SVM+ technique might exhibit better performance than the sentiment classification model that applies the SVM technique. However, the following problems were observed in this experiment. The first problem is related to the selection of group data. In this experiment, the year of product issue was selected as group data in the same experimental environment. Because the groups did not have a close relationship with each other in the sentiment classification model that applies the SVM+ technique, this model exhibited the lowest prediction performance among other models. The second problem is related to the amount of data.

This study conducted an experiment by extracting sample data and employing a small amount of data. The results obtained from the experiments indicate that the model that applies the SVM+MTL technique exhibited the best performance, while the model that applies the SVM+ technique exhibited a better prediction performance than the model that applies the SVM technique. However, as the amount of data increased, the model that applies the SVM+

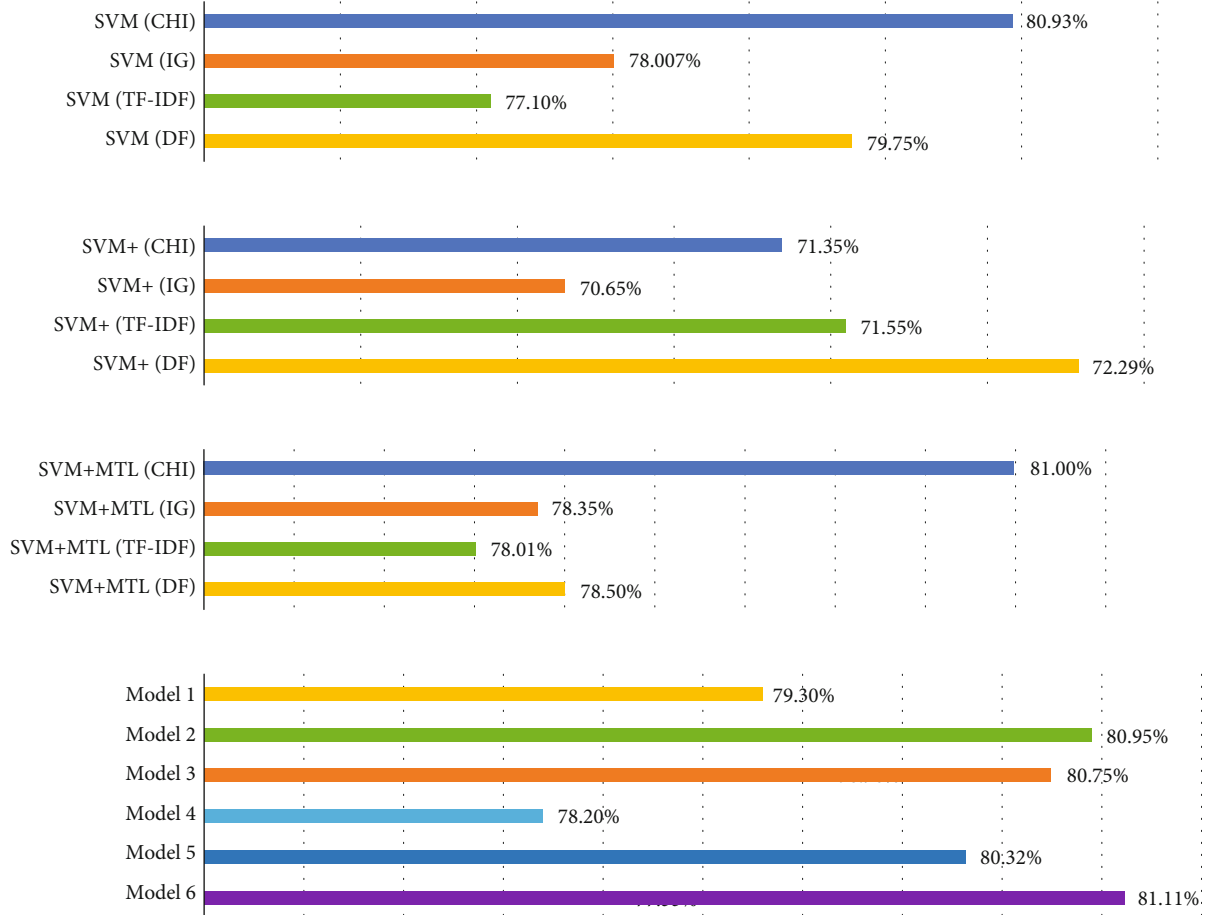


FIGURE 14: Comparison of overall model performance in book reviews.

TABLE 4: Experiment results of sentiment classification models in four cases.

	Document frequency			TF-IDF			Information gain			Chi-square statistic		
	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL	SVM	SVM+	SVM+MTL
Movie												
Accuracy	76.95	71.15	77.15	75.41	67.33	75.54	77.65	67.38	77.91	84.11	75.92	82.32
F-measure	77.18	75.52	72.37	72.99	66.43	73.49	76.80	66.78	77.18	75.52	75.24	81.78
Game												
Accuracy	80.27	75.27	79.55	77.56	69.60	77.86	79.80	73.58	80.10	80.27	75.22	80.72
F-measure	81.06	75.03	78.45	78.62	68.61	78.79	80.97	73.29	81.37	82.06	74.93	81.38
Music												
Accuracy	75.37	73.58	75.27	72.29	75.23	72.04	75.37	73.58	75.27	80.45	72.39	79.85
F-measure	75.06	75.37	75.03	72.14	74.93	71.37	55.06	73.73	75.03	81.88	73.30	79.63
Book												
Accuracy	79.75	72.29	78.50	77.10	71.55	78.01	78.00	70.65	78.35	80.93	71.35	81.00
F-measure	77.48	70.99	78.28	75.40	67.30	75.50	77.44	70.15	79.93	78.42	70.90	79.55

technique exhibited the poorest performance, contrary to expectations. The final problem is related to the selection of the kernel function and parameter adjustment. Accordingly, the results obtained from the experiments conducted in this study indicated that an appropriate kernel function and parameter should be selected.

In this study, the performance of the proposed sentiment classification model was evaluated by comparing it with the prediction performance of single models. The results obtained from analyzing the performance of sentiment classification models based on movie, game, music, and book reviews indicate that the sentiment classification model that

TABLE 5: Experiment results of document-level sentiment classification models in four cases.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Movie						
Accuracy	84.83	85.68	84.23	83.78	85.68	82.81
F-measure	84.09	85.61	82.09	82.05	85.59	82.10
Game						
Accuracy	81.37	82.05	80.88	79.95	83.25	81.04
F-measure	80.23	81.29	79.48	79.27	82.33	80.38
Music						
Accuracy	79.09	77.55	78.96	81.49	76.63	77.35
F-measure	78.46	76.23	75.79	80.96	75.73	76.65
Book						
Accuracy	79.30	80.95	80.75	78.20	80.32	81.11
F-measure	77.42	78.90	78.55	77.44	70.15	78.93

applies the chi-square statistic exhibited the best performance among the term information extraction techniques compared.

5. Conclusion

This study established a document-level term-based sentiment classification model by collecting user reviews generated in an e-commerce environment and using reviews that contained subjective opinions of users. In addition, this study analyzed the proposed model according to four cases to increase efficiency and information power, perform effective term extraction, and present the comparative performance of sentiment classification models. Accordingly, an experiment was conducted to verify the prediction performances of single models and the proposed model, respectively.

Specifically, this study developed sentiment classification models that apply SVM+ and SVM+MTL techniques and compared the prediction performance of these models with the sentiment classification model that applies the SVM technique, which is frequently used in existing studies. The aforementioned techniques are frequently applied in research to solve binary classification problems. This study is distinguished from existing studies in that it applied these techniques to opinion mining research. Furthermore, it proposed an integrated sentiment classification model that addressed the limitations of single models and exhibited enhanced prediction performance, which was verified in experimental results.

Furthermore, this study applied the proposed method to an opinion mining research, which requires the analysis of a considerable amount of data, and verified that the proposed model effectively analyzed reviews. The proposed model identified sentimental expressions of existing users toward products in an e-commerce environment based on various POS. Accordingly, it is deduced that research on various POS can be adopted as an effective methodology for analyzing decisions to be made by potential users in the process of purchasing products. The proposed model has advantages because it can collect, use, and monitor information on certain products, sur-

veys, or industries from users on social media in actual time, to reflect analytical results to customers. It is expected that the proposed model can contribute to increasing the spread of opinion mining for sentiment classification based on user reviews. In addition, this model can also improve e-commerce services and the quality of these services by enabling users to obtain desired information more easily and purchase or use desired products based on various types of recommendation information provided to them.

Further research should be conducted to combine studies at the document level with those at the sentence level. As this experiment did not reflect the latest user reviews updated, the flow of positive or negative tendencies of opinions over time was not accurately predicted. This study also has a limitation in that highly influential terms adopted as input variables cannot represent documents. Considering this limitation, follow-up research should be conducted to form a sentiment dictionary at the sentence level, apply various input variables to the document-level sentiment classification model, and ultimately increase prediction performance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) Funded by the Ministry of Education (Nos. NRF-2021R1I1A3049061 and NRF-2017R1A6A1A03015496).

References

- [1] M. Kang, J. Ahn, and K. Lee, "Opinion mining using ensemble text hidden Markov models for text classification," *Expert Systems with Applications*, vol. 94, pp. 218–227, 2018.
- [2] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information fusion*, vol. 36, pp. 10–25, 2017.
- [3] S. B. Marzizarani and H. Sajedi, "Opinion mining with reviews summarization based on clustering," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1299–1310, 2020.
- [4] M. Husnain, M. M. S. Missen, N. Akhtar, M. Coustaty, S. Mumtaz, and V. S. Prasath, "A systematic study on the role of SentiWordNet in opinion mining," *Frontiers of Computer Science*, vol. 15, no. 4, pp. 1–19, 2021.
- [5] P. Sánchez-Núñez, M. J. Cobo, C. De Las Heras-Pedrosa, J. I. Peláez, and E. Herrera-Viedma, "Opinion mining, sentiment analysis and emotion understanding in advertising: a bibliometric analysis," *IEEE Access*, vol. 8, pp. 134563–134576, 2020.
- [6] F. Alattar and K. Shaalan, "A survey on opinion reason mining and interpreting sentiment variations," *IEEE Access*, vol. 9, pp. 39636–39655, 2021.
- [7] A. Kumar, A. S. Veerubhotla, V. T. Narapareddy, V. Aruru, L. B. Neti, and A. Malapati, "Aspect term extraction for opinion mining using a Hierarchical Self-Attention Network," *Neurocomputing*, vol. 465, pp. 195–204, 2021.
- [8] J. Yu, J. Jiang, and R. Xia, "Global inference for aspect and opinion terms co-extraction based on multi-task neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 1, pp. 168–177, 2019.
- [9] F. Ali, K. S. Kwak, and Y. G. Kim, "Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification," *Applied Soft Computing*, vol. 47, pp. 235–250, 2016.
- [10] Y. Cai, K. Yang, D. Huang et al., "A hybrid model for opinion mining based on domain sentiment dictionary," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2131–2142, 2019.
- [11] C. Hung and S. J. Chen, "Word sense disambiguation based sentiment lexicons for sentiment classification," *Knowledge-Based Systems*, vol. 110, pp. 224–232, 2016.
- [12] H. Ha, H. Han, S. Mun, S. Bae, J. Lee, and K. Lee, "Visualization of movie recommendation system using the sentimental vocabulary distribution map," *Journal of The Korea Society of Computer and Information*, vol. 21, no. 5, pp. 19–29, 2016.
- [13] M. H. Arif, J. Li, M. Iqbal, and K. Liu, "Sentiment analysis and spam detection in short informal text using learning classifier systems," *Soft Computing*, vol. 22, no. 21, pp. 7281–7291, 2018.
- [14] D. H. Pham and A. C. Le, "Learning multiple layers of knowledge representation for aspect based sentiment analysis," *Data & Knowledge Engineering*, vol. 114, pp. 26–39, 2018.
- [15] N. F. Da Silva, E. R. Hruschka, and E. R. Hruschka Jr., "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, vol. 66, pp. 170–179, 2014.
- [16] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.
- [17] D. Bi, J. Kong, X. Zhang, and J. Yang, "Analysis on health information acquisition of social network users by opinion mining: case analysis based on the discussion on COVID-19 vaccinations," *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [18] M. Keyvanpour, Z. K. Zandian, and M. Heidarypanah, "OMLML: a helpful opinion mining method based on lexicon and machine learning in social networks," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–17, 2020.
- [19] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014.
- [20] A. Tripathy, A. Anand, and S. K. Rath, "Document-level sentiment classification using hybrid machine learning approach," *Knowledge and Information Systems*, vol. 53, no. 3, pp. 805–831, 2017.
- [21] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013.
- [22] P. D. Turney and M. L. Littman, "Measuring praise and criticism," *ACM Transactions on Information Systems (tois)*, vol. 21, no. 4, pp. 315–346, 2003.
- [23] T. Y. Kim, S. B. Pan, and S. H. Kim, "Sentiment digitization modeling for recommendation system," *Sustainability*, vol. 12, no. 12, p. 5191, 2020.
- [24] N. F. Ibrahim and X. Wang, "A text analytics approach for online retailing service improvement: evidence from Twitter," *Decision Support Systems*, vol. 121, pp. 37–50, 2019.
- [25] F. Hu, L. Li, X. Xu, J. Wang, and J. Zhang, "Opinion extraction by distinguishing term dependencies and digging deep text features," *Neural Computing and Applications*, vol. 31, no. 9, pp. 5419–5429, 2019.
- [26] A. Thakkar and K. Chaudhari, "Predicting stock trend using an integrated term frequency-inverse document frequency-based feature weight matrix with neural networks," *Applied Soft Computing*, vol. 96, article 106684, 2020.
- [27] N. S. M. Nafis and S. Awang, "An enhanced hybrid feature selection technique using term frequency-inverse document frequency and support vector machine-recursive feature elimination for sentiment classification," *IEEE Access*, vol. 9, pp. 52177–52192, 2021.
- [28] E. O. Omuya, G. O. Okeyo, and M. W. Kimwele, "Feature selection for classification using principal component analysis and information gain," *Expert Systems with Applications*, vol. 174, article 114765, 2021.
- [29] Y. Liu, X. Yi, R. Chen, Z. Zhai, and J. Gu, "Feature extraction based on information gain and sequential pattern for English question classification," *IET Software*, vol. 12, no. 6, pp. 520–526, 2018.
- [30] C. Z. Liu, Y. X. Sheng, Z. Q. Wei, and Y. Q. Yang, "Research of text classification based on improved TF-IDF algorithm," in *2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE)*, pp. 218–222, Lanzhou, China, 2018.
- [31] Z. Zhu, J. Liang, D. Li, H. Yu, and G. Liu, "Hot topic detection based on a refined TF-IDF algorithm," *IEEE access*, vol. 7, pp. 26996–27007, 2019.
- [32] C. Iwendi, S. Ponnar, R. Munirathinam, K. Srinivasan, and C. Y. Chang, "An efficient and unique TF/IDF algorithmic model-based data analysis for handling applications with big data streaming," *Electronics*, vol. 8, no. 11, p. 1331, 2019.
- [33] A. Jalilifard, V. F. Caridá, A. F. Mansano, R. S. Cristo, and F. P. C. da Fonseca, "Semantic sensitive TF-IDF to determine word relevance in documents," in *Advances in Computing and Network Communications*, pp. 327–337, Springer, Singapore, 2021.

- [34] M. Labani, P. Moradi, F. Ahmadizar, and M. Jalili, "A novel multivariate filter method for feature selection in text classification problems," *Engineering Applications of Artificial Intelligence*, vol. 70, pp. 25–37, 2018.
- [35] P. Sur, Y. Chen, and E. J. Candès, "The likelihood ratio test in high-dimensional logistic regression is asymptotically a rescaled chi-square," *Probability Theory and Related Fields*, vol. 175, no. 1-2, pp. 487–558, 2019.
- [36] V. Vapnik, *Estimation of dependencies based on empirical data*, Springer-Verlag, 2nd edition, 2006.
- [37] A. Altan and S. Karasu, "The effect of kernel values in support vector machine to forecasting performance of financial time series," *The Journal of Cognitive Systems*, vol. 4, no. 1, pp. 17–21, 2018.
- [38] J. Xu, C. Xu, B. Zou, Y. Y. Tang, J. Peng, and X. You, "New incremental learning algorithm with support vector machines," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 11, pp. 2230–2241, 2019.
- [39] T. Y. Kim, S. H. Bae, and Y. E. An, "Design of smart home implementation within IoT natural language interface," *IEEE Access*, vol. 8, pp. 84929–84949, 2020.
- [40] M. Lapin, M. Hein, and B. Schiele, "Learning using privileged information: SVM+ and weighted SVM," *Neural Networks*, vol. 53, pp. 95–108, 2014.
- [41] B. Ribeiro, C. Silva, A. Vieira, A. Gaspar-Cunha, and J. C. das Neves, "Financial distress model prediction using SVM+," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, Barcelona, Spain, 2010.
- [42] C. Serra-Toro, V. J. Traver, and F. Pla, "Exploring some practical issues of SVM+: is really _privileged_ information that helps?," *Pattern Recognition Letters*, vol. 42, pp. 40–46, 2014.
- [43] L. Liang, F. Cai, and V. Cherkassky, "Predictive learning with structured (grouped) data," *Neural Networks*, vol. 22, no. 5-6, pp. 766–773, 2009.