JOURNAL OF BASIC SCIENCE AND ENGINEERING

RANDOM SAMPLING DATA PREPROCESSED MUTUAL INFORMATION ROCCHIO SENTIMENT CLASSIFICATION BASED SENTIMENT ANALYSIS FOR RECOMMENDATION SYSTEM

Dr. P.Suresh [1], G.Shumalatha [2]

[1]Associate Professor and HOD, Dept of Computer Science, Salem Sowdeswari College – Govt Aided, Salem, Tamilnadu.
[2] Research Scholar, Bharathiyar University, Coimbatore, Tamilnadu.

DOI: https://doie.org/10.0905/Jbse.2024105685

Abstract

Sentiment Analysis is considered as one of significant problem to be resolved as itdetermines the opinions and emotions of the users through written contents. A Random Sampling Data Preprocessed Mutual Information Rocchio Sentiment Classification (RSDPMIRSC) Technique is proposed for providing efficient recommendation to an item with higher accuracy and lesser time. Initially, customer review and feedback of an item are collected from the large database. After that, the collected customer reviews are preprocessed in RSDPMIRSC Technique using term based random sampling process for automatically detecting and removing the stop word from the reviews. The designed process functioned through iterating the selected customer review at random manner. After removing the stop words from reviews, classification process is carried out in RSDPMIRSC Technique using Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm to perform the customer review classification as positive and negative with higher accuracy and lesser time. MIRSCis a supervised method used to compute mutual information between given term and classes (i.e., positive and negative). The designed models provided the information how much the term related to the particular class. Based on the customer review classification, the recommendation regarding particular item purchase is given to the user. Experimental evaluation of RSDPMIRSC Technique is carried out on factors such as time complexity, classification accuracy, error rate and preprocessing time with respect to number of customer reviews.

Keywords: Customer reviews, Mutual Information Rocchio Sentiment Classification, Recommendation, Sentimental analysis, Term based random sampling

1. Introduction

Sentiment analysis is the process of categorizing the opinions in text to discover whether attitude towards particular product. Due to large amount of reviews, sentiment analysis plays an essential role to extract the useful information of reviews. Sentiment classification is a difficult task in sentiment analysis. Recommendation system determines the rating or preference which user gives to an item.

Recommender systems use any filtering method, namely collaborative filtering and content-based filtering. Collaborative filtering constructs the model from user past behavior and similar decisions made by other users. Collaborative filtering is used to predict the ratings for items. But, the collaborative filtering approaches suffer from three issues, namely cold start, scalability, and sparsity. In addition, accurate recommendation was not provided to the user. In order to address these problems, a novel technique is introduced in this research work for sentiment classification.

A hybrid context-aware recommendation framework named SocialRec was designed in [1] with a rating inference for personalized recommendations of various items. However, recommendation performance was poor. A Support Vector Machines (SVM) classifier model was employed in [2]with application of two feature selection techniques to identify positive, neutral, and negative emotions. But, classification accuracy was not at required level.

The sentiment-specific word embedding's were collected in [3] from Arabic tweets and utilized in Arabic Twitter sentiment classification. But, the computational overhead was not reduced. A three-way enhanced convolutional neural network model termed 3W-CNN was introduced in [4]. But, the complexity level was not reduced using 3W-CNN.

A new neural network model termed AttDR-2DCNN was introduced in [5] with two parts. However, the preprocessing time was not reduced using AttDR-2DCNN modeling. A new stock recommendation system was introduced in [6] for finding the correlation between Guba-based sentiment of retail investors and the stock market trends in China. But, the recommendation was not provided in accurate manner by using stock recommendation system.

A sentiment classification model was introduced in [7] for addressing the domain-sensitive and data imbalance problems during the sentiment classification. However, the classification time was not reduced by using sentiment classification model. A domain attention model was introduced in [8] for multi-domain sentiment analysis. But, the feature extraction accuracy was not improved.

An optimization model was designed in [9] for finding relationship concealed in the content features. A new multi-aspect user-interest model was introduced in [10] for recommender systems to enhance the recommendation and prediction accuracy. But, the time complexity was not reduced by using new multi-aspect user-interest model.

For addressing the above mentioned conventional issues in sentimental analysis for recommendation system, RSDPMIRSC Technique is developed. The main contributions of RSDPMIRSC Technique is described in below,

- ✓ To improve the classification performance of sentimental analysis, by using RSDPMIRSC Technique is designed to provide better recommendation system when compared to conventional methods. On the contrary to traditional works, proposed RSDPMIRSC Technique is intended with the help of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithms.

- ✓ To efficiently carry out preprocessing task with minimal amount of time complexity

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

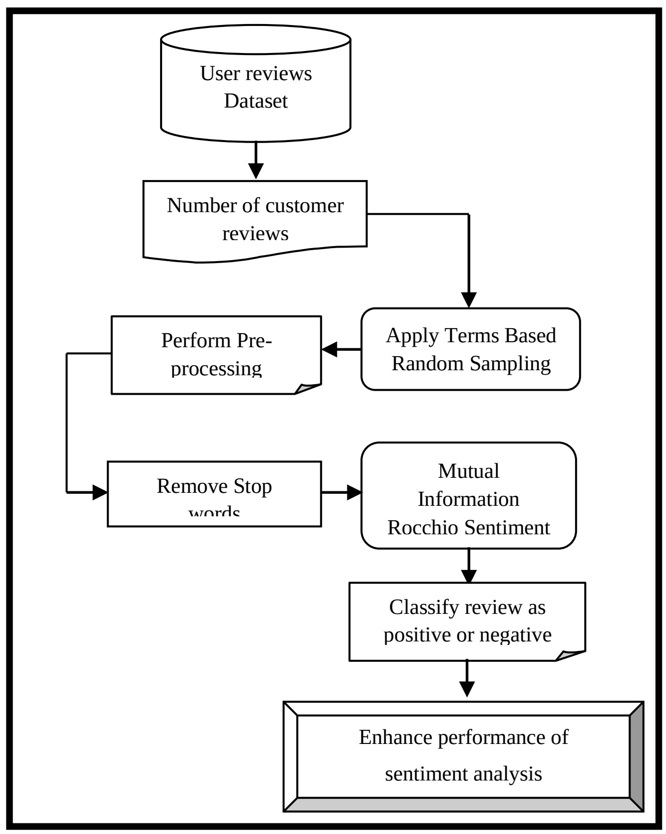
during the sentiment classification process when compared to existing works, T-RS algorithm is developed in proposed RSDPMIRSC Technique. On the contrary to state-of-the-art works, T-RS algorithm allocates random weights to each term in a customer review from sentiword Net dictionary for eliminating irrelevant terms.

- ✓ To achieve better sentiment classification performance with a lower time when compared to conventional works, MIRSC algorithm is introduced in proposed RSDPMIRSC Technique. On the contrary tostate-of-the-art works,MIRSC algorithm employs mutual information measurement for efficient classification of reviews as positive or negative sentiment with higher accuracy. From that, MIRSCalgorithm attains enhanced recommendation performance when compared to existing works.

The rest of the paper is designed as follows; In Section 2, RSDPMIRSC Technique is explained with assist of architecture diagram. In Section 3, simulation settings are presented and the performance result of RSDPMIRSC Technique is analyzed in Section 4. Section 5 shows the literature survey. Section 6 shows the conclusion of the paper.

2. Random Sampling Data Preprocessed Mutual Information Rocchio Sentiment Classification Technique

A recommender system provides suggestions to users, in multiple contexts. The aim of a recommender system is to give recommendations to users regarding items or products. In conventional works, few research works have been designed for performing sentimental analysis with help of different classification algorithms. However, misclassification of customer reviews using existing work was higher which reduces the accuracy of sentiment analysis. Besides, the amount of time required for sentiment analysis is also not minimized. To addresses the above mentioned issues, Random Sampling Data Preprocessed Mutual Information Rocchio Sentiment Classification (RSDPMIRSC) Technique is developed. On the contrary to state-of-the-art works, RSDPMIRSC Technique is proposed by combining Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm. The architecture diagram of RSDPMIRSC Technique is depicted in below Figure 1.



Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Figure 1 Architecture Diagram of RSDPMIRSC Technique for Sentiment Analysis

Figure 1 demonstrates the overall flow process ofRSDPMIRSC Technique to attain better sentiment analysis and recommendation system. As shown in the above figure, RSDPMIRSC Technique initially takes user reviews dataset i.e.OpinRank Review Datasetasinput. Then, RSDPMIRSC Technique applies Terms Based Random Sampling (T-RS) method to carry out preprocessing. During this process, T-RS method take outs the stop words in each input customer reviews with lower amount of time consumption. Subsequently, RSDPMIRSC Technique appliesMutual Information Rocchio Sentiment Classification (MIRSC) algorithm where it classify customer reviews as positive or negative sentiment with higher accuracy. From that, RSDPMIRSC Technique improves sentiment analysis performance to recommend purchase items to the users.

2.1 Terms Based Random Sampling Process

The Terms Based Random Sampling (T-RS) algorithm is designed with aim of increasing the performance of preprocessing task with minimal time for effective sentimental analysis. The T-RS algorithm is one of the types of Biased Sampling technique. Thisrepresents each and every terms in the whole dataset has the different probability of being selected. In this T-RS, the terms are selected based on its weight.

On the contrary to conventional works, T-RS algorithm is proposed by using the random term weighting concepts. In T-RS algorithm, random Term weighting is the assignment of random numerical values to terms that represent their importance. Besides to that, random Term weighting considers the relative importance of individual words because not all the terms in a given reviews are of equal importance.The termswith higher weight value are selected as significant terms for sentiment analysis process. The terms with lower weight value are removed as stop words or irrelevant in order to minimize the time complexity of sentiment analysis during the recommendation process. The process involved in Terms Based Random Sampling is shown in below Figure 2.

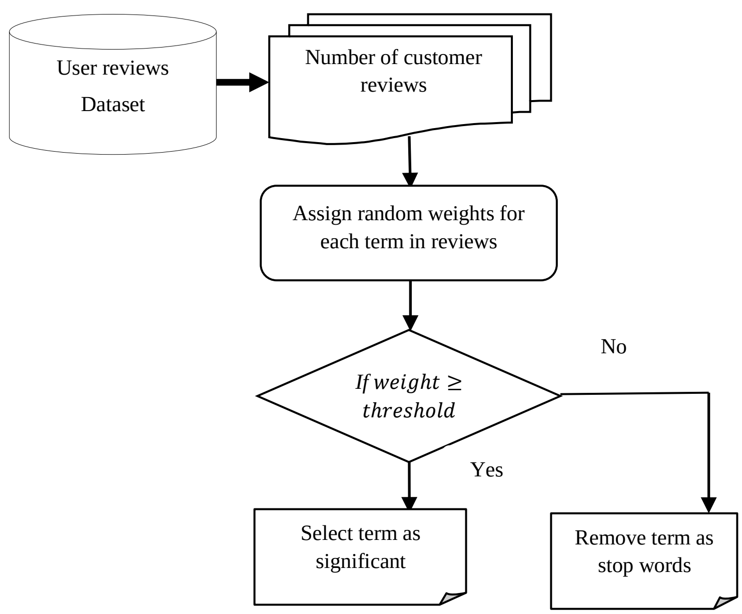


Figure 2 Terms Based Random Sampling for preprocessing

Figure 2 demonstrates the block diagram of T-RS algorithm for efficiently accomplishing preprocessing during the sentiment classification. As shown in the above figure, T-RS algorithm at the beginning obtains userreviews dataset i.e. OpinRank

Review Dataset. The number of customer reviews in input dataset 'DS' is mathematically represented as follows,

$$DS = \beta_1, \beta_2, ..\beta_n \quad (1)$$

From the above equation (1), 'β_n' denotes the number of customer reviews in given dataset. Followed by for each terms 'T_i' in customer review, then random term weighting is mathematically performed by using sentiwordNet dictionary as follows,

$$w_i = \sum_{T_i=1}^{n} Randam\ weight\ () \quad (2)$$

From the above equation (2), 'w_i' represent random weight given for each term from sentiwordNet dictionary. In T-RS algorithm, Senti Word Net is an opinion lexicon obtained from the WordNet database where each term is associated with numerical score. Next, T-RS algorithm defines threshold weight value. Then T-RS algorithm compares weight value of each term with threshold weight value and thereby significantly selects the terms with maximum weight value as significant for accurate sentiment classification. Also, T-RS algorithm eliminates the terms with lower weight value as unrelated during the sentiment analysis process. The algorithmic processes of Terms Based Random Sampling is explained as follows,

// Terms Based Random Sampling Algorithm
Input: User reviews dataset 'DS', number of customer reviews '$\beta_1, \beta_2, ..\beta_n$'
Output: Eliminate irrelevant words with minimal preprocessing time
Step 1:Begin
Step 2:For each input customer reviews 'β_i'
Step 3: **For** each terms 'T_i'
Step 4: Randomly assigns weights from dictionary using (2)
Step 5: **If** $(weight \geq threshold)$, **then**
Step 6: Choose term as important for sentiment analysis
Step 7: **Else**
Step 8: Remove term as stop words
Step 9: **End if**
Step 10: **End For**
Step 11: **End For**
Step 12:End

Algorithm 1 Terms Based Random Sampling

Algorithm 1 shows the step by step process of Terms Based Random Sampling. As demonstrated in the above algorithmic steps, T-RS model at first takes number of customer reviews as input. For each term in customer reviews, then T-RS model randomly gives weights with help of dictionary. Followed by, T-RS model checks if weight value of term is greater than predefined threshold. If the above condition is true, T-RS model finds the terms as a significant in order to efficiently perform sentiment classification. Otherwise, T-RS model eliminate the terms as stop words. From that, T-RS model reduce the amount of time taken for sentiment analysis when compared to state-of-the-art works.

2.2 Mutual Information Rocchio Sentiment Classification Model

The Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm is designed with intention of enhancing the classification accuracy of sentiment analysis. The MIRSC algorithm is based on Vector Space Model. MIRSC algorithm is proposed by mutual information measurement in the conventional Rocchio classification algorithm. The designed MIRSC algorithm is easy to implement and also very fast learner. The main idea of MIRSC algorithm is to represent each terms 'T_i' in input reviews as a vector in a vector space so that customer

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

reviews with similar features have similar vectors. On the contrary to state-of-the-art classification algorithms, finally designed MIRSC algorithm assigns a class label to each customer reviews based on estimated mutual information. Thus, MIRSC algorithm efficiently performs the sentiment analysis process with higher accuracy and minimal time as compared to existing works. The process involved in MIRSC algorithm is shown in below Figure 3.

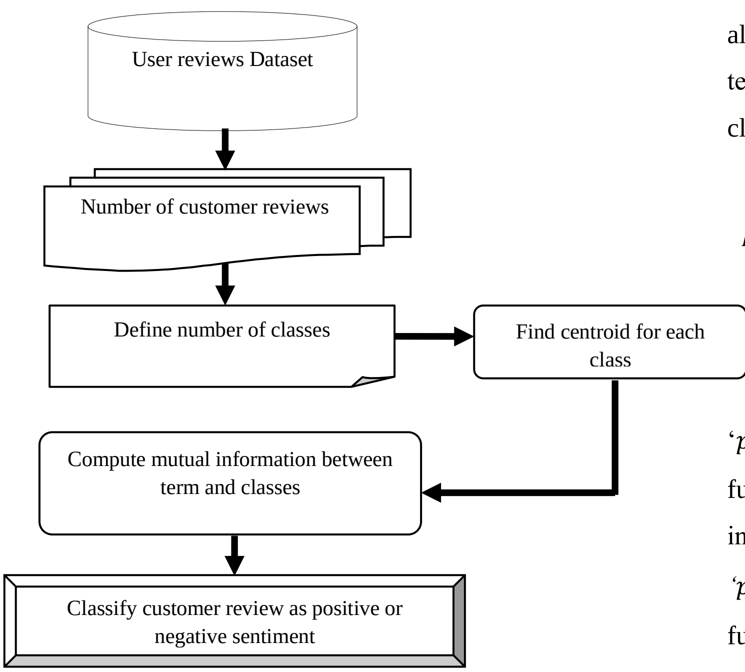


Figure 3 MIRSC algorithm Process

Figure 3 presents the flow processes of the MIRSC algorithm to get better accuracy for sentiment analysis. As illustrated in the above figure, the MIRSC algorithm at first takes the number of customer reviews as input. Then, MIRSC algorithm initializes the number of classes. In the MIRSC algorithm, a number of classes are defined as two namely positive sentiment class and negative sentiment class. After initializing the classes, MIRSC algorithm identify the centroid 'α' for each class 'c_j' using below,

$$\alpha_{(c_j)} = \frac{1}{|n_c|} \sum_{\beta_i \epsilon DS} v(\beta_i) \tag{3}$$

From (3) '$\alpha_{(c_j)}$' represent the centroid of 'i^{th}' class whereas 'n' refers the number of customer reviews in given dataset whose class is 'c'. Here, '$v(\beta_i)$' indicates the normalized vector of customer reviews 'β_i'. The centroid is determined as the center of mass of its class members. Subsequently, MIRSC algorithm estimates mutual dependence between each terms of input customer review 'β_i' and class'c_j'using below,

$$MI(\beta_i, c_j) = \int_{\beta_i} \int_{c_j} pr(\beta_i, c_j) \log \frac{pr(\beta_i, c_j)}{pr(\beta_i) pr(c_j)} d\beta_i dc_j \tag{4}$$

From above mathematical expression (4), '$pr(\beta_i, c_j)$' indicates the joint probability density function between the two between each terms of input customer review 'β_i' and class 'c_j'. Here, '$pr(\beta_i)$' and '$pr(c_j)$' represents the marginal density functions. The mutual dependence determines how similar the joint distribution '$pr(\beta_i, c_j)$' is to the products of the factored marginal distribution. In MIRSC algorithm, mutual dependence '$MI(\beta_i, c_j)$' yields values in range of '0' to '$+\infty$'. According to the measured mutual dependence value, MIRSC algorithm classifies each input customer review as positive or negativesentiment with improved accuracy. This process ofMIRSC algorithm is repeated until all the customer reviews in given datasetare classified into a particular class. The algorithmic process of MIRSC algorithm is explained in below.

// **Mutual Information Rocchio Sentiment Classification Algorithm**
Input: number of customer reviews '$\beta_1, \beta_2, .. \beta_n$'
Output: Enhanced classification accuracy for sentiment analysis
Step 1: Begin
Step 2: Initialize number of class'c_i'
Step 3: Determine centroid '$\alpha_{(c_i)}$' for each class using (3)
Step 4: **For** each customer reviews 'β_i'
Step 5: Calculate mutual dependence value using (4)
Step 6: Classify the review as positive or negative sentiment
Step 7: **End For**
Step 8:End

Algorithm 2Mutual Information Rocchio Sentiment Classification

Algorithm 2 explains the step by step process of MIRSC algorithm. As shown in the above algorithmic steps, the MIRSC algorithm initializes number of customer reviews as input. Then, MIRSC algorithm initializes the number of class's. Next, MIRSC algorithm computes the centroid for each class and consequently estimates the mutual information between each customer review and class centroid. Finally, the MIRSC algorithm assign a class label to the customer review based on determined mutual information value. Thus, the MIRSC algorithm attains better classification performance to accurately find the customer opinions about different products with a lower amount of time. Therefore, the RSDPMIRSC Technique increases the accuracy of sentiment analysiswhen compared to conventional methods.

3. EXPERIMENTAL SETTINGS

To evaluate the performance, both the proposed RSDPMIRSC Technique and conventional two methods are implemented in Java language by using OpinRank Review Dataset. The input OpinRank Review Dataset [21] is collected from UCI machine learning repository from http://archive.ics.uci.edu/ml/datasets/opinrank+review+dataset. This dataset includes user reviews of cars and hotels gathered from Tripadvisor and Edmunds. They constitute full reviews from Tripadvisor (~259,000 reviews) and Edmunds (~42,230 reviews), From that, proposed RSDPMIRSC Technique considers a different number of car customer reviews in the range of 50-500 for model-years 2007, 2008, and 2009 to conduct the experimental process. This dataset contains 140-250 cars for each model year with attributes such as dates, author names, favorites and the full textual review. The RSDPMIRSC Technique conducts experimental work by using parameters such as preprocessing time, classification accuracy, and error rate and time complexity. The performance result of RSDPMIRSC Technique is compared against conventional hybrid context-aware recommendation framework called SocialRec [1] and Support Vector Machines (SVM) classifier [2].

4. RESULTS

In this section, the experimental result of RSDPMIRSC Technique is discussed. The performance of RSDPMIRSC Technique is compared against existing hybrid context-aware recommendation framework called SocialRec [1] and Support Vector Machines (SVM) classifier [2] respectively. The performance of RSDPMIRSC

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Technique is analyzed along with the following metrics with the help of tables and graphs.

No. of customer reviews	Preprocessing Time (ms)		
	RSDPMIRSC	SocialRec	SVM classifier
50	11	13	15
100	13	14	17
150	15	17	18
200	16	18	19
250	18	19	20
300	20	22	23
350	21	23	25
400	22	25	26
450	23	26	27
500	25	28	30

4.1 Preprocessing Time

In RSDPMIRSC Technique, Preprocessing Time (PT) estimates amount of time utilized to remove the stop words in input customer reviews. The preprocessing time is mathematically determined as follows,

$$PT = Number of customer\ reviews * time\ (RSSR) \quad (5)$$

From the above mathematical expression (5), preprocessing time is determined in the terms of milliseconds (ms) whereas '$time\ (RSSR)$' denotes the amount of time taken for removing stop words in single customer review.

Table 1 Tabulation result ofPreprocessing Time for Sentiment analysis

Table 1 demonstrates performance result of preprocessing time is determined during the process of sentiment analysis and recommendation system with respect to dissimilar number of customer reviews in the range of 50 to 500 using three techniques. From the above gets tabulation results, it is expressive that the preprocessing time using proposed RSDPMIRSC Technique is lower to achieve better sentiment classification performance with lower amount of time when compared to traditional hybrid context-aware recommendation framework called SocialRec [1] and Support Vector Machines (SVM) classifier [2]. Based on the above tabulation values, the graph is drawn in below for comparative analysis of proposed and existing works.

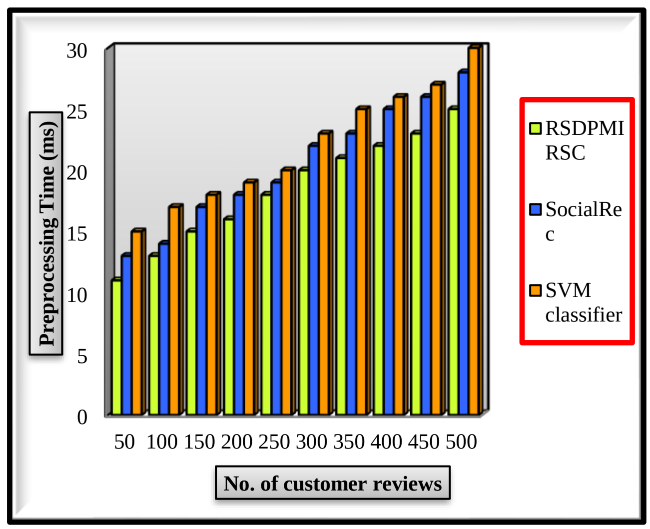


Figure 4 Experimental Result of Preprocessing Time versus Number of Customer Reviews

Figure 4 presents graphical result analysis of preprocessing time according to various number of customer reviews in the range of 50 to 500 using three techniques namely proposed RSDPMIRSC Technique and conventional SocialRec [1] and SVM classifier [2]. As demonstrated in the above graphical figure, proposed RSDPMIRSC Technique gives lower preprocessing time during the process of sentimental analysis for better recommendation system when compared to state-of-the-art SocialRec

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

No. of customer reviews	Classification accuracy (%)		
	RSDPMIRSC	SocialRec	SVM classifier
50	88	72	78
100	90	79	81
150	87	78	80
200	89	81	82
250	88	74	78
300	87	78	81
350	85	78	81
400	85	77	79
450	84	75	78
500	83	74	77

[1] and SVM classifier [2]. This is owing to application of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm in proposed RSDPMIRSC Technique on the contrary to existing works. This supports for proposed RSDPMIRSC Technique to decrease the amount of time consumed to take away the stop words in input customer reviews when compared to other works SocialRec [1] and SVM classifier [2]. Thus, proposed RSDPMIRSC Technique minimizes preprocessing time of sentiment analysis by 10 % and 17 % when compared to state-of-the-art SocialRec [1] and SVM classifier [2] respectively.

4.2 Classification Accuracy

In RSDPMIRSC Technique, Classification Accuracy (CA) is measured as the ratio of number of customer reviewsthat are exactly classified to the total number of customer reviews. The classification accuracy is mathematically determined as follows,

$$CA = \frac{P_{cc}}{n} * 100 \quad (6)$$

From the above mathematical equation (6), 'P_{cc}' indicates number of customer reviews correctly classified and 'n' refers total number of customer reviewstaken for carried outing the experimental process. The classification accuracy of sentiment analysis and recommendation system is evaluated in terms of percentage (%).

Table 2 Tabulation result of Classification Accuracy for Sentiment analysis

Table 2 present's experimental result of classification accuracy is measured during the process of sentiment analysis and recommendation system based on varied number of customer reviews in the range of 50 to 500 using three techniques. From the above tabulation results, it is clear that the classification accuracy using proposed RSDPMIRSC Technique is very higher to suggest best purchase items to users when compared to state-of-the-art SocialRec [1] and SVM classifier [2]. Depends on the above tabulation value, the graph is plotted in below for performance analysis of proposed and existing works.

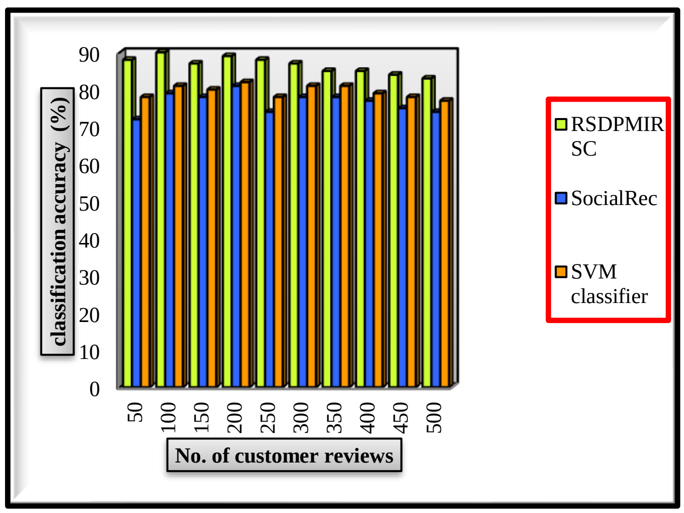


Figure 5 Experimental Result of Classification Accuracy versus Number of Customer Reviews

Figure 5depictsperformance result analysis of classification accuracy with respect to different number of customer reviews in the range of 50 to 500 using three techniques namely proposed RSDPMIRSC Technique and traditional SocialRec [1] and SVM classifier [2]. As shown in the above graphical representation, proposed RSDPMIRSC Technique provides enhanced accuracy during the process of sentimental analysis to achieve better recommendation performance when compared to conventional SocialRec [1] and SVM classifier [2]. This is because of application of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm in proposed RSDPMIRSC Technique on the contrary to state-of-the-art works. This helps for proposed RSDPMIRSC Technique to improve the ratio of number of customer reviews that are correctly classified when compared to other works SocialRec [1] and SVM classifier [2]. As a result, proposed RSDPMIRSC Technique increases classification accuracy of sentiment analysis by 13 % and 9 % when compared to traditional SocialRec [1] and SVM classifier [2] respectively.

4.3 Error Rate

In RSDPMIRSC Technique, the Error Rate (ER) computes the ratio of a number of customer reviews wrongly classified to the total number of customer reviews. The error rate is mathematically calculated as follows,

$$ER = \frac{P_{ic}}{n} * 100 \qquad (7)$$

From the above mathematical formula (7), 'P_{ic}' signifies number of customer reviews in correctly classified and 'n' refers total number of customer reviews. The error rate of sentiment analysis is determined in terms of percentage (%).

Table 3 Tabulation result of Error Rate for Sentiment analysis

No. of customer reviews	Error rate (%)		
	RSDPMIRSC	SocialRec	SVM classifier
50	12	28	22
100	10	21	19
150	13	22	20
200	11	19	18
250	12	26	22
300	13	22	19
350	15	22	19
400	15	23	21
450	16	25	22
500	17	26	23

Table 3showstabulation result analysis of error rate is calculated during the process of sentiment classification and recommendation system depends on different number of customer reviews in the range of 50 to 500 using three techniques. From the experimental evaluation results, it is significant that the error rate using proposed RSDPMIRSC Technique is lower to precisely classify reviews as positive sentiment or negative sentiment when compared to existing SocialRec [1] and SVM classifier [2]. According to the above tabulation value, the graph is designed in below for relative analysis of proposed and conventional works.

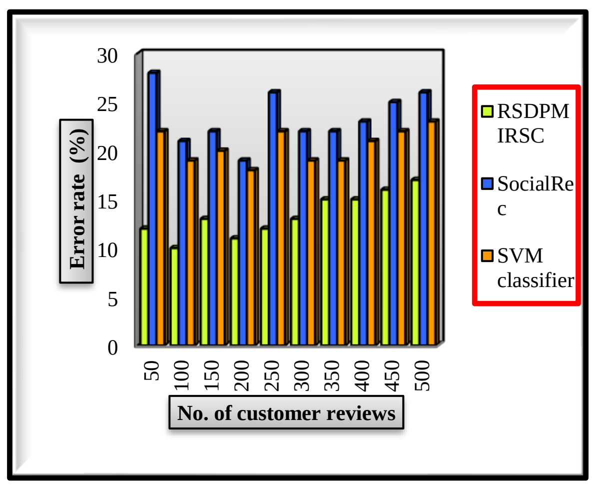


Figure 6 Experimental Result of Error Rate versus Number of Customer Reviews

Figure 6depictscomparative result analysis of error rate based on various number of customer reviews in the range of 50 to 500 using three techniques namely proposed RSDPMIRSC Technique and existing SocialRec [1] and SVM classifier [2]. As illustrated in the above graphical depiction, proposed RSDPMIRSC Technique attains lower error rate for accurate sentimental classification and thereby enhancing the recommendation performance when compared to traditional SocialRec [1] and SVM classifier [2]. This is due to application of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm in proposed RSDPMIRSC Technique on the contrary to conventional works. These aids for proposed RSDPMIRSC Technique to reduce the ratio of number of customer reviews that are inaccurately classified when compared to other works SocialRec [1] and SVM classifier [2]. Accordingly, proposed RSDPMIRSC Technique minimizes error rate of sentiment analysis and recommendation system by 42 % and 35 % when compared to existing SocialRec [1] and SVM classifier [2] respectively.

4.4 Time Complexity

In RSDPMIRSC Technique, Time Complexity (TC) calculates the amount of time taken for classifying input customer reviews into corresponding class i.e. positive sentiment or negative sentiment. The time complexity is mathematically estimated using below,

$$\text{TC} = ET - ST \quad (8)$$

From the above mathematical equation (8), 'ST'represents a starting time and 'ET' signifies an ending time of sentiment classification. The time complexity is determined in terms of milliseconds (ms).

Table 4 Tabulation result of Time Complexity for Sentiment analysis

No. of customer reviews	Time Complexity (ms)		
	RSDPMIRSC	SocialRec	SVM classifier
50	19	26	29
100	22	28	32
150	25	32	35
200	28	35	38
250	30	39	41
300	33	42	44
350	35	46	47
400	38	49	50
450	40	53	55
500	43	56	58

Table 4 illustrates comparative result of time complexity is estimated during the process of sentiment analysis and recommendation system along with diverse number of customer reviews in the range of 50 to 500 using three techniques. From the above acquired experimental results, it is considerable that the time complexity using proposed RSDPMIRSC Technique is very minimal to accurately categorize reviews as positive sentiment or negative sentiment and thereby recommending best products to users when compared to conventional SocialRec [1] and SVM classifier [2]. By using the above tabulation a value, the graph is intended in below for analyzing the proposed and existing performances.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

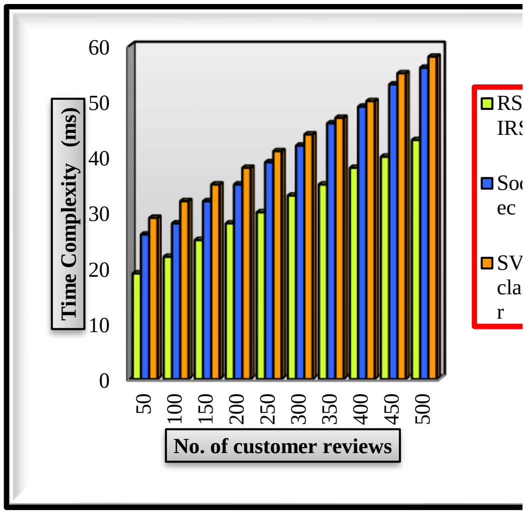


Figure 7 Experimental Result of Time Complexity versus Number of Customer Reviews

Figure 7portraysexperimental result analysis of time complexity with respect to different number of customer reviews in the range of 50 to 500 using three techniques namely proposed RSDPMIRSC Technique and state-of-the-art SocialRec [1] and SVM classifier [2]. As depicted in the above graphical illustration, proposed RSDPMIRSC Technique gets minimal time complexity for effective sentimental analysis when compared to conventional SocialRec [1] and SVM classifier [2]. This is because of application of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm in proposed RSDPMIRSC Technique on the contrary to traditional works. This assists for proposed RSDPMIRSC Technique to diminish the amount of time taken for categorizing input customer reviews into consequent class i.e. positive sentiment or negative sentiment when compared to other works SocialRec [1] and SVM classifier [2]. Hence, proposed RSDPMIRSC Technique reduces time complexity of sentiment analysis and recommendation system by 23 % and 28 % when compared to conventional SocialRec [1] and SVM classifier [2] respectively.

5. RELATED WORKS

Capsule network was constructed in [11] for performing the sentiment analysis in domain adaptation scenario with the semantic rules (CapsuleDAR). However, the error rate was not minimized. A deep learning-based method was introduced in [12] with unified feature set based on representative of word embedding, sentiment knowledge, sentiment shifter rules, statistical and linguistic knowledge. But, time complexity was not reduced using deep learning-based method.

A gated recurrent neural network with sentimental relations (GRNN-SR) was introduced in [13] to gather sentimental relative information. But, the classification accuracy was not improved using GRNN-SR. A sophisticated algorithm depending on deep learning was introduced in [14] for classification. But, the computational cost was not minimized using sophisticated algorithm.

A Feature-Based Fusion Adversarial Recurrent Neural Networks (FARNN-Att) was introduced [15]. A regularization method was employed to increase robustness and generalization ability. A sentiment analysis technology was studied in [16] with deep neural network to explore reviews, insert ratings to reviews. But, the error rate was not minimized using sentiment analysis technology.

A deep learning model was introduced in [17] to process user comments and to generate the user rating for user recommendations. However, the accuracy level was not improved using deep learning model. A knowledge-based recommendation system (KBRS) was employed in [18] with application of

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

deep learning for increasing the accuracy of sentiment analysis. But, computational time was more.

A recommendation system that employs aspect-based opinion mining (ABOM) was designed in [19] with help of deep learning technique to get better the accuracy. However, the ratio of number of customer reviews that are properly classified was lower. A user recommendation algorithm for social networks was introduced in [20] depends on sentiment analysis and matrix factorization. But, accuracy during recommendations process was minimal.

6. CONCLUSION

The RSDPMIRSC Technique is intended with the objective of increasing the performance of recommendation system via accurate classification of customer reviews. The aim of RSDPMIRSC Technique is attained with the application of Terms Based Random Sampling (T-RS) and Mutual Information Rocchio Sentiment Classification (MIRSC) algorithm on the contrary to conventional works. The proposed RSDPMIRSC Technique diminishes the amount of time required to remove the unrelated words in input each customer reviews when compared to traditional works. Also, the proposed RSDPMIRSC Technique enhances the ratio of number of customer reviews that are exactly classified as positive sentiment or negative sentiment to design effective recommendation system when compared to conventional works. Besides, proposed RSDPMIRSC Technique reduces the ratio of number of customer reviews that are mistakenly classified when compared to state-of-the-art works. In addition, proposed RSDPMIRSC Technique decreases the amount of time taken for classifying input customer reviews when compared to conventional works. The simulation result depicts that RSDPMIRSC Technique presents better recommendation system performance with an enhancement of classification accuracy and minimization of preprocessing time for sentiment analysis when compared to state-of-the-art works.

REFERENCES

[1] Rizwana Irfan, Osman Khalid, Muhammad Usman Shahid Khan, Faisal Rehman, Atta Ur Rehman Kha, “SocialRec: A Context-Aware Recommendation Framework With Explicit Sentiment Analysis”, IEEE Access, Volume 7, Pages 116295 – 116308, August 2019

[2]Mohammad A.Hassonah, RizikAl-Sayyed, AliRodanabAla’ M.Al-Zoubi, Ibrahim Aljarah, Hossam Faris, “An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on Twitter”, Knowledge-Based Systems, Elsevier, Pages 1-40, December 2019

[3] Nora Al-Twairesh and Hadeel Al-Negheimish, “Surface and Deep Features Ensemble for Sentiment Analysis of Arabic Tweets”, IEEE Access, Volume 7, Pages 84122 – 84131,June 2019

[4] Yuebing Zhang, Zhifei Zhang, Duoqian Miao and Jiaqi Wang, “Three-way Enhanced Convolutional Neural Networks for Sentence-level Sentiment Classification”, Information Sciences, Elsevier, Volume 477, Pages 55-64,March 2019

[5] Fagui Liu, Jingzhong Zheng, Lailei Zheng and Cheng Chen, “Combining attention-based bidirectional gated recurrent neural network and two-dimensional convolutional neural network for document-level sentiment classification”,

Neurocomputing, Elsevier, Pages 1-12, September 2019

[6] Yunchuan Sun, Mengting Fang and Xinyu Wang, “A novel stock recommendation system using Guba sentiment analysis”, Personal and Ubiquitous Computing, Springer, Volume 22, Issue 3, Pages 575–587, June 2018

[7] Yijing Li, Haixiang Guo, Qingpeng Zhang, Mingyun Gu, Jianying Yang, “Imbalanced text sentiment classification using universal and domain-specific knowledge”, Knowledge-Based Systems, Elsevier, Volume 160, Pages 1-15, November 2018

[8] Zhigang Yuan, Sixing Wu, Fangzhao Wu, Junxin Liu and Yongfeng Huang, “Domain Attention Model for Multi-Domain Sentiment Classification”, Knowledge-Based Systems, Volume 155, Pages 1-10, September 2018

[9] Haijun Zhang, Yanfang Sun, Mingbo Zhao, Tommy W. S. Chow and Q. M. Jonathan Wu, “Bridging User Interest to Item Content for Recommender Systems: An Optimization Model”, IEEE Transactions on Cybernetics, Pages 1-13, March 2019

[10] Lihua Sun, Junpeng Guo and Yanlin Zhu, “A multi-aspect user-interest model based on sentiment analysis and uncertainty theory for recommender systems”, Electronic Commerce Research, Springer, Pages 1–26,September 2018

[11] Bowen Zhang, Xiaofei Xu, Min Yang, Xiaojun Chen and Yunming Ye, “Cross-Domain Sentiment Classification by Capsule Network with Semantic Rules”, IEEE Access, Volume 6, Pages 58284 – 58294, October 2018

[12] Asad Abdi, Siti Mariyam Shamsuddin, Shafaatunnur Hasan and Jalil Piran, “Deep learning-based sentiment classification of evaluative text based on Multi-feature fusion”, Information Processing and Management, Elsevier, Volume 56, Pages 1245-1259, 2019

[13] Chaotao Chen, Run Zhuo and Jiangtao Ren, “Gated recurrent neural network with sentimental relations for sentiment classification”, Information Sciences, Elsevier, Volume 502, Pages 268–278, 2019

[14] Meng Wang, Zhen-Hu Ning, Chuangbai Xiao and Tong Li, “Sentiment Classification Based on Information Geometry and Deep Belief Networks”, IEEE Access, Volume 6, Pages 35206 – 35213,June 2018

[15] Yaohong Ma, Hong Fan, and Cheng Zhao, “Feature-Based Fusion Adversarial Recurrent Neural Networks for Text Sentiment Classification”, IEEE Access, Volume 7, Pages 132542 – 132551, September 2019

[16] Hyeon-woo An and Nammee Moon, “Design of recommendation system for tourist spot using sentiment analysis based on CNN-LSTM”, Journal of Ambient Intelligence and Humanized Computing, Springer, Pages 1-11,September 2019

[17] Rung-Ching Chen and Hendry, “User Rating Classification via Deep Belief Network Learning and Sentiment Analysis”, IEEE Transactions on Computational Social Systems, Volume 6, Issue 3, Pages 535 – 546, June 2019

[18] Renata Lopes Rosa, Gisele Maria Schwartz, Wilson Vicente Ruggiero, Demóstenes Zegarra

Rodrígue, “A Knowledge-Based Recommendation System That Includes Sentiment Analysis and Deep Learning”, IEEE Transactions on Industrial Informatics, Volume 15 , Issue 4, Pages 2124 – 2135, April 2019

[19]AminuDa'u, NaomieSalim, IdrisRabiu, AkramOsman, “Recommendation system exploiting aspect-based opinion mining with deep learning method”, Information Sciences, Elsevier, Volume 512, Pages 1279-1292, February 2020

[20]ChongchaoCai, HuahuXu, “A topic sentiment based method for friend recommendation in online social networks via matrix factorization”, Journal of Visual Communication and Image Representation, Elsevier, Volume 65, Pages 1-5, December 2019

[21]OpinRank Review Dataset:http://archive.ics.uci.edu/ml/datasets/opinrank+review+dataset