

Dimensionality Reduction In Sentiment Analysis Using Colony–Support Vector Machine

Harjeet Kaur, Prabhjeet Kaur

Abstract: With the advent of technology in different areas, sentiment analysis is an emerging field with advancement of internet of things. Analyzing the sentiments, views, estimations, conduct and feeling through textual or written format is known as sentiment analysis or opinion mining. Generally, huge amount of data is available on the internet for example the data that is present in blogs, assessment websites, feedback forums and so forth. Internet of things is incentive to these developments. A large amount of information that is available on internet is amorphous and manageable from internal areas in websites, evaluation sites, and review forums. Presently, a number of people prefer online shopping, because there are several sources for buying products; thus making it less time consuming and cost effective. In addition, sentiments, views and feelings of the customers in the reviews and comments can be categorized as positive, negative or neutral; that helps the new customers to make decision about the quality of the product and about the company. In this paper, a unique approach is built on a specific subject by trusting the reviews on social sites. Proposed approach contains a list of the words that is used to design information based training group (positive keyword and negative keyword). In the research work Colony-SVM (Support vector machine) is used for classification of sentiments and KPCA to reduce feature dimensionality. Performance analysis has defined some likely consequences on the existing work when compared to our proposed model. Along with this, feature vector method is implemented using two stages after pre-processing method. In this research, initially data is collected from the social sites like Amazon etc along with extraction of unique features from gathered information then adding it to features vector and value set. In this research work step by step description of reviews and sentiment analysis is performed. This paper also defines a comparative analysis of Machine learning algorithms with accuracy rate, class precision and recall.

Index Terms: Sentiment Analysis, Opinion mining Machine learning, Internet of things.

I. INTRODUCTION

With the advancement of the blogging, review sites and social networking sites, huge amount of the data has been shared among the individuals in last few decades. A number of individual describe their opinion with set of features. The main feature those work in is giving the feedback to the companies about the products and list of the customers. The

process of the selection of the large amount of the data is the sentiment analysis [1]. The recovery of the decision about the various products and classification is done on the basis of the recommendation or positive or may be negative feedback. The feedback regarding the products can be categorized as positive, negative or neutral [2]. Sentiment analysis is calculation of the decision, emotional view, occurrence, characteristics and feedback of the persons [3]. The opinion mining also known as sentiment analysis is the process where the opinions play the main role for making the choices and recognizing the other person's choice. Hence, there are no computational approaches required for making the choice of the emotion. In last few decades, opinion from the friends and the families was asked for building the choices. But in making the choices about the products and the services by the organizations group of survey has been done. Hence, the advancement in the content in the social sites which results vast transformation in the world. In current world of technology, the customers express the views on the blogs and social networking sites. The process is the sentiment analysis through websites [4][5]. At present, the reviews are seen on the web and discovering sentiment behind the customer reviews done by the industries and companies. The survey is done for gathering the choices of the customers about the products. In different case, the sentiment analysis is hidden in post on the blogs. In the era of the internet technology, there is wide range of the advancement in the internet communication. The web data is not organized in the structured format so it is not easy to detect the data automatically [6]. The automatic detection of the behavior (feeling, judgment, sentiments and beliefs) of the human is the sentiment analysis. In addition, classifying the information of sentiment analysis into positive; negative and other emotion pattern recognizes the sentimental views of the data [2]. The high amount of the data taken for the different opportunities and one who gets the data may have unreliable data. The method used for the computation of the same emotional behavior is sentiment analysis [7]. The recognition of the behavior of the individual who pens or speaks for unique topic is the method of the text sentiment analysis [10]. In the social media analysis, there is an advancement of the audios and videos this method is text analysis of the sentiment data [11]. The method of the automatic recognition of the emotion of the person in the form of the texts is called as the sentiment analysis; this method is also called as the opinion mining method [12].

Revised Manuscript Received on June 15, 2019

Harjeet Kaur, Dept. of CSE, Sachdeva Engineering College For Girls, Gharuan, Mohali, India.

Prabhjeet Kaur, Dept. of CSE, Sachdeva Engineering College For Girls, Gharuan, Mohali, India.

At Present, recognition and the distribution of the thoughts and opinions are in the form of the texts. With the advancement of the multimedia, there are different methods of determining the views, thoughts through internet with the increment in the content data [8]. The method of the sentiment analysis is the identification and the extraction of the subjective data using the natural language and analysis of the text data [9]. Sentiment analysis is the method of the evaluation of the decision of the sentiments of the author in one domain or the multi domain. The main Applications of Sentiment Analysis are:

- i) In Business and Financial domain.
- ii) Senti-WordNet as lexical resource.
- iii) Socialites.
- iv) Voting advice applications.
- v) Monitoring the opinion of the people.

In existing work, an automatic analysis of such information field of SA is used. SA data is subjected to several data pre-processing approaches and then verifying the idea in the comments and classifying them. Open Source Data Tool analysis simulation tool is known as RM (Rapid Miner) is used to evaluate performance and step-by-step description of review processing. In base paper also defines analysis of the comparative study of various methods such as Naïve Bayes (NB) and SVM (Support Vector Machine).

In the proposed approach, kernel Principal Component analysis is used by extracting the features in vector form to reduce the dimensions in the reviews. After that, the novel colony support vector machine is implemented and then using feature extraction method that helps in the selection of the extracted features and classification is done in the form of the binary format (0, 1).

This paper contains 5 sections that describe sentiment analysis and Machine Learning Methods. Section I contains Introduction about the sentiment analysis, its applications, Existing work and proposed work. Section II presents various literature surveys in the field of sentiment or opinion mining. Section III and IV elaborates reading the proposed methodology and Evaluation of research work. The Last section V defines conclusion and future scope of sentiment analysis.

II. RELATED WORK

Ren, R. and Wu, D. D et al., 2018[13] proposed a research on the analysis of the sentiments by using machine learning approach with support vector machine. In this research, the analysis of the day of week along with investor sentiment is taken into consideration which results in more realistic and valid sentiment indexes. In this research, the method of the integration of the support vector machine is through a realistic window. The experimental results describe about the probable approach about the features of the stock market and forecasting behavior of the data. This also implies that sentiments will have very important information about resource intrinsic value. The method of the searching of the values regarded as the indication of the stock market data.

Hoogendoorn, M., and Berger, T et al., 2017[14] studied about the prediction of the models by the recognition the writing of the patients that suffers from social anxiety. In this research, the extraction of the features of the health of the

patients is done by taking the text written by the patients into consideration, Also their style of writing and the basic sentiment of the messages is also taken into account. In this research, the prediction models used is produced on the basis of the machine learning algorithm based on the data sets of the 69 patients. The outcome of the therapy is described using area under the curve (AUC) and the result of the area that may be under the curve were based on the precision of about 0.78% on the entire treatment period. In this research, the prediction of the emotion of the patients and the prediction of the modeling is done on the basis of the anxiety symptoms. The analyzer may analyses the prediction models on the basis of the symptoms on weekly basis. The data sets of the 69 patients are taken for the diagnosis on the weekly basis. The outcomes were taken on the basis of the measurement of the social phobia by the therapist in three stages. Firstly, the demographic information is taken and after that the two months treatment was done by the therapist. Finally, the prediction of the models is done from the received emails over the given period of the time.

Tai, C. H., Tan, Z. H and Chang, Y. S et al., 2016[15] demonstrated a research on the target and the strength of the emotion of the person through views from the social network. The recognition of the positive emotions of the peoples was done through the online approaches. The detection was quite difficult due to overloading of the data. Hence, in this research, an emotion recognition system was established on the basis of the latent allocation method, sentiment word recognition method. By using this method the emotion and the objective of the feeling of the person was detected. The experimental research is done by the collection of posts of an individual on his/her social network site which indicates an individual's feelings and intensity behind the feeling with the help of feeling distinguisher system. In this research, the recognizing of the emotions is classified as strong positive, strong negative, positive and negative.

Scharl, A., Herring, D and, Rafelsberger, W et al., 2017[16] investigated the semantic approaches about the views of the emotions. In this research, the extraction of the context data was done using the cognitive therapy method. The visualization method includes were philological, interpersonal and spatial context approach. The research describes about the recognition of the visual approach in real time data of the multiple coordinated views for recognizing the contextual approach via semantic dimensions. The proper recognition of the context data describes about the extracted features. In this research, the contextual data recognize the mutual data between the objects that include location and organization and social media which are Twitter, Facebook, Google+ and YouTube.

Sindhu C et al., 2017 [17] studied about the presence of huge amount of data in e-commerce sites. In order to control the study of such data Sentiment Analysis is used. An open source data tool Rapid miner is used to perform opinion mining. SVM (Support Vector Machine) is used to classify opinions into different classes that are positive, negative and neutral. Comparative study of algorithms like SVM and Naïve Bayes also discussed in this research work.



Dhanalakshmi et al., (2016) [18] Investigated sentiment analysis on student feedback using supervised machine learning algorithms. The main aim of this study was to find the duality of student feedback. The study utilizes blend of machine learning technique and natural language processing techniques on the input information assembled from module assessment study aftereffects of Centre East School, Oman. The data collected is subjected to various pre processing steps which are then used to prepare the calculations for binomial arrangement. This paper additionally introduces a relative investigation on performance of various algorithms like SVM, NB, KNN (K Nearest Neighbor) .It is concluded in the study that SVM works best for data set.

Guellil et al., (2015) [19] investigated that sentiment analysis is a growing concept and attracting many researchers. In order to address the emergence of social media as there is a huge amount of data that is misused to identify hidden knowledge present in the data. In this paper study of 60 social media mining papers was proposed. Among these 60 papers 60% of papers were published between 2013 and 2015.This paper also addresses various issues on sentiment analysis for example using different languages may lead to many irresolvable problems. It was also found that Sentiment analysis using supervised learning algorithms may gave best results but require lot of time and effort. By addressing the various issues it was concluded that solving these issues will result in better future work.

Claypo, N., et al., (2015) [20] proposed strategy to do opinion mining on millions of Thai restaurant reviews. This study concluded that it is hard to do opinion mining using unsupervised techniques therefore it was proposed to do opinion mining on Thai restaurant review using K-Means clustering and MRF feature selection. K-Means clustering is used to categorize negative and positive reviews and MRF is used to reduce dimensionality in the feature set, which make it computationally fast in time. What’s more, K-means can accomplish the best grouping presentation, when contrasted and Self-Sorting out Guide, Fuzzy C-Means.

III. PROPOSED RESEARCH WORK

Sentiment Analysis uses C-SVM (Colony-Support Vector Machine) classification tool in order to classify the sentiments as Negative, Positive and Neutral. The following research methodology was implemented in this research work

A. Standard Dataset

Search the dataset from the online sites such as Amazon, Snap deal and Flip kart. These are open source dataset. The information is collected through the reviews written by the users on various social media sites. These reviews may contain nouns, pronouns, adjectives etc and extraction is used to segment individual sentences thus we use this information to create our own dataset containing positive and negative keywords. The comments on the various methods and materials for example an electronic product and its corresponding rating are also defined in the different instances.

In sentiment analysis main problem is un-structured data and then using unsupervised machine learning methods to

classify the un-structured data with the help of various pre-processing methods.

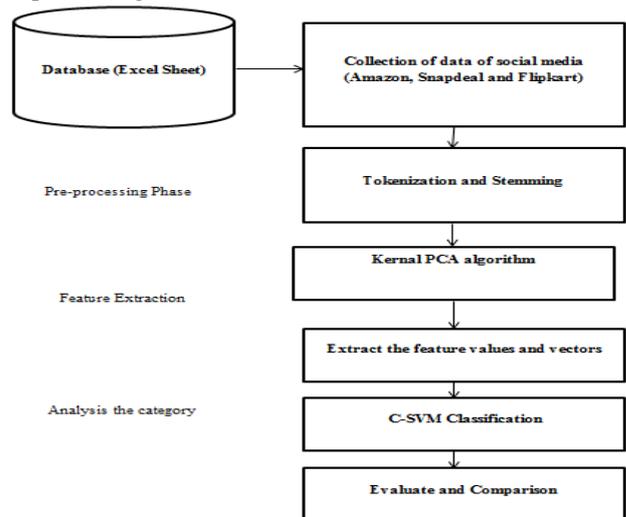


Fig. 1 Flow Chart of the Proposed Work

Above figure 1 describes the flowchart of the proposed work. It shows how the research work was implemented.

In this research work we have created our own Knowledgebase. The Knowledgebase contains two classes for the dataset: Negative Keywords and Positive Keywords. These are some of the example of the negative and positive keywords present in our knowledgebase.

“NEGATIVE KEYWORDS”

'didn't '
'bad'
'donot'
'abnormal'
'abolish'
'abominable'
'absence'
'absurdity'
'accost'
'acerbic'
'acrimonious'
'addicting'
'admonishingly'
'aimless'
'backache'
'not'
'ugly'

“POSITIVE KEYWORDS”

'excellent service'
'excellent Phone'
'achieve'
'efficient'
'Good'
'awesome'
'confidence '
'complement'
'satisfied'
'original product '
'great value'
'Happy'

'intelligence'
'hello'
'beautiful'
'gorgeous'
'handsome'

B. Data pre-processing

In this phase, the input data is pre-processed before extracting the characteristics that includes removing the tags, eliminating missing instances, removing missing instances or attributes with the another variables in-order to neglect the errors. Dataset used in EXCEL or CSV file is used and comments are expressed in ENGLISH language.

Main Pre-processing steps are:

- (i) Tokenization
- (ii) Stemming

The tokenization process is the act of division of series of string into segments like words, keywords, phrases and signs elements known as tokens. It can be individual phrases, words or complete sentences. In this procedure of tokenization, few alphabets such as punctuation marks are rejected.

It is utilized in CS (Computer Science), where tokenization plays a main role and also in huge section in the process of LA (Lexical Analysis).

Example: He is a good Seller.

Token Output:

	1
1	He
2	is
3	a
4	good
5	Seller.
6	

Fig. 2 Tokenization process

Stemming process is defined as the process of optimizing word to its word stem or root forms normally an inscribed written word format. It is not required to be equal stem, even if this stem isn't in itself a reasonable root.

Example: She is singing a song.

Stem Output: Sing.

C. Kernel PCA used for Feature Extraction

The feature extraction method is developed to fetch the unique properties in the vector form like matrix (r,c). KPCA (Kernel Principal Component Analysis) method is used to extract the individual components in the form of Eigen values and vector values. It is an extension to the PCA algorithm. It computes data into a novel Feature space that helps in dimensionality reduction. It captures the overall variance of designs and helps in removing noises in the images.

Steps of Feature Extraction (KPCA) Method [21].

Let $k = 1$.

For each data point $x_{ii} \in X_{trr}$; if $y_{ii}^{lr} = k$ otherwise let $Z_{im}^{lr} = +1$ otherwise let $Z_{im}^{lr} = -1$.

Next, for assigning the nearest method, use kernel regression ($Z_{jm}^{ll}, J=1, \dots, m_u$) to the testing set, using new labels $Z_{ir}^{lm} = 1, \dots, l_{mtii}$ as input for labeling method.

If $k < q$, then increase k by 1, and then return to step 2, otherwise continue to next step.

For each data point $y_t \in Y^{ll}$. Iterate through classes $k=1, \dots, q$.

For first k for which $Z_{jkk}^{ll} = 1$, let $Z_{jj}^{ll} = k$ and go to next data point. If no such k exists, then let $y_{jj}^{ll} = q$.

D. Selection of the feature (C-SVM) Algorithm

The classify matrix of similarity values is created and given to machine learning algorithms that is colony-SVM for its training. The colony-SVM classification method selects the kernel feature set; divide the valuable data to calculate the minimum and maximum distance of the feature vector. If equal distance found out in the divided phase the modification phase will help to transfer the equal data. Training of Colony-SVM is performed in MATLAB (2016a). Training and Testing is done with 70-30 ratio. 70 % part of data is given to training and 30% part of data is given to testing. Presentation of each classifier is analyzed; evaluation is done on basis of presentation parameters those are true positive, true negative, accuracy rate, recall rate and precision.

Proposed algorithm (C-SVM algorithm)

```

Start round
do while (Ending criteria not-correct) – cycle loop
    do until (each ant or comments completes a round) – round loop
        Local trail configure
    End do
    Study round
    Global round update
End do
    
```

It is used to resolve the minimum rate issues. It may normally have N number of comments and un-directed arcs. There are twice working number of comments for the nodes or reviews: either backwards or forward.

The comments memory gives them information to re-trace the route it has followed while searching for the sink node or review.

Before, moving backward on their memorized route, they change any loops from it. While moving rearward, the comments consent pheromones on the arcs they crossed.

An initialization of the search procedure, constant amount of pheromone is allotted to all comments.



When positioned at a comment I ant k uses the pheromone trail to evaluate the probability of selecting j as next_node.

$$p_{ij}^k = \begin{cases} \tau_{ij}^k \\ \sum N_i^k \tau_{ij}^k \end{cases} \text{ if } j \in N_i^k, \text{ if } j \notin N_i^k \dots\dots\dots (i)$$

Where, N_i^k is the closest of comment k when in node of words i .

When the word (i,j) is crossed, the pheromone val modifies as follows:-

$$\tau_{ij} < - - \tau_{ij} + \Delta\tau^k \dots\dots\dots (ii)$$

Using this rule, the probability increases that forth-coming reviews will use this word.

After that word k has changed to next phase, the pheromones evaporate by the following eq at all the words.

$$\tau_{ij} < - - (C - p)\tau_{ij}, \forall (i, j) \in A \dots\dots\dots (iii)$$

X, Y defined with training labelled data, $\alpha < 0$,

$\alpha < \text{partial_train_SVM}$.

$C < \text{some value}$.

Repeat

For all $\{x_i, y_i\}, \{x_j, y_j\}$ do

Optimize α and α_j do,

End for

Until no_changes in α or other resource constraint criteria met.

End

E. Evaluation

In this section, description about result and performance analysis of the sentiment analysis is mentioned. In this research work, performance of the applied proposed model (C-SVM) is checked in order to verify the accuracy rate and class precision. The un-supervised methods that were used in SVM and ACO are combined in this Research work.

Table 1. Test Review analysis

Test Review	Sentiment Analysis
He is a good Seller.	Positive
She is very Beautiful	Positive
She is bad Girl	Negative
He is good student but his handwriting is bad	Neutral
It might be good	Neutral

Table 1 describes the sentiment behind the review tested in the given knowledge base. The Knowledge base will categorize the reviews into 3 categories as shown in table 1 after determining its sentiment. In this testing phase to analyses the review is based on distance comparison between training set and testing set feature in this dataset. If distance same then result evaluation is done on the basis of precision and accuracy rate.

Table 2 contains the Experimental results of our research work. As shown in Table 2 it is clear that this table compares the performance of the proposed colony-SVM with the SVM. The results shows that precision, recall and the true positive rate of the C-SVM is higher than the traditional SVM, which implies that the proposed model is more feasible.

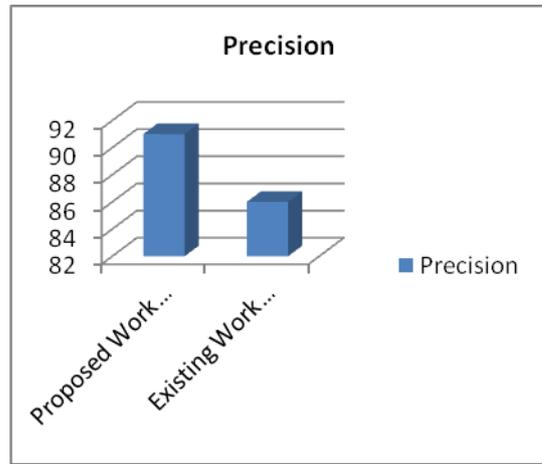


Fig. 3 Comparison – Precision

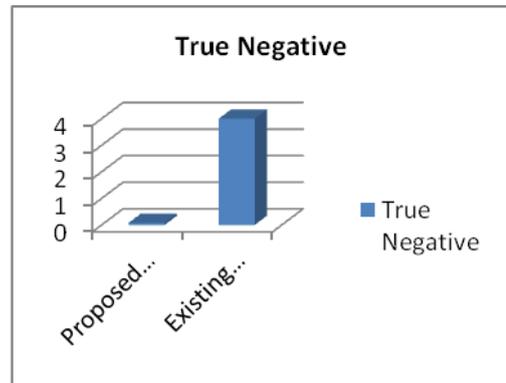


Fig. 4 Comparison – True Negative rate

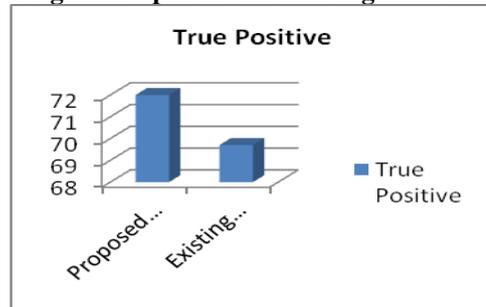


Fig. 5 Comparison- True Positive rate

Table 2. Performance Analysis of proposed work

Parameter Metrics	Proposed Work (C-SVM)	Existing Work (SVM)
Precision	91	86
True Negative	0.08	4
True Positive	72	69.7
Recall	98	96.1



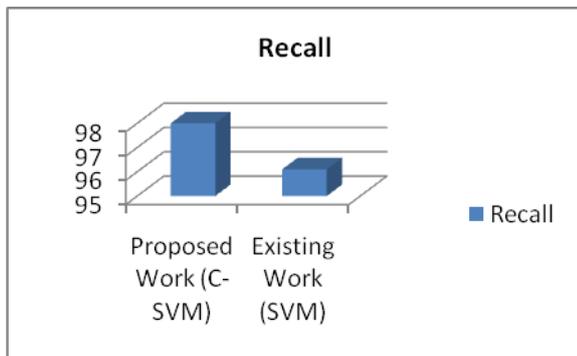


Fig. 6 Comparison- Recall

In this section comparison of performance parameters of the two classifiers using SVM and C-SVM algorithm were evaluated. Performance parameters used in the proposed model that are Precision, True negative rate, True positive rate and Recall rate has been compared with the traditional SVM in figure 3,4,5,6 respectively.

In this paper, classifiers were analyzed using MATLAB 2016a Simulation. In research work two types of normal classifiers were used (SVM and C-SVM) for sentiment classification. SVM and C-SVM Classifiers are designed using MATLAB construct in methods. Classifier is designed using GUI (Graphical User Interface). Performances of these classifiers are defined in Figure 3, 4, 5 and 6. All of this classifier has almost similar performance metrics. The main issue is main tokenization process. Language like in British and American are defined to be un-segmented words. For tokenization process these languages need some addition lexical information. Errors in Stemming process means dissimilar stems are stemmed to the similar roots called True Negative and True Positive Rate.

IV. CONCLUSION AND FUTURE SCOPE

This research work it is concluded that SA (Sentiment Analysis) on the various comments, reviews given by the users have been evaluated using Machine Learning methods that are designed using C-SVM model. The novel approach classifies the text into 3 categories that is POSITIVE; NEGATIVE AND NEUTRAL depends on their distance similarity. SA has been established efficiently to predict consumer feelings by studying SOCIAL DATA and REVIEWS. In this work, we implement a new approach that helps in extracting the consumer review about a specific subject by relying on SOCIAL MEDIA AND SHOPPING SITES's comments and reviews which are expressed by the users time to time. This paper combines the advantages of KPCA and C-SVM. First the data is collected from the user that is the reviews an individual post on various social media sites then this collected data is subjected to various pre-processing phases such as feature extraction and feature selection with help of KPCA and C-SVM is used as final text classifier. The Experimental results as shown in Table 2 indicates that performance of C-SVM is better than that of SVM which makes our research work more feasible and effective. The Future work is for implementing Sentiment Analysis by using methods like Decision Tree and K-nearest Neighbor Algorithm. To improve the data pre-processes

word embedding DNNs (Deep Neural Network) can also be used in addition to the C-SVM.

REFERENCES

- Poria, S., Peng, H., Hussain, A., Howard, N., & Cambria, E. (2017). Ensemble application of convolutional neural networks and multiple kernel learning for multimodal sentiment analysis. *Neurocomputing*, 261, 217-230.
- Tan, S., & Zhang, J. (2008). An empirical study of sentiment analysis for chinese documents. *Expert Systems with applications*, 34(4), 2622-2629.
- Benamara, F., Cesarano, C., Picariello, A., Recupero, D. R., & Subrahmanian, V. S. (2007, March). Sentiment analysis: Adjectives and adverbs are better than adjectives alone. In *ICWSM* (pp. 1-7).
- Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27, 16-32.
- Alessia, D., Ferri, F., Grifoni, P., & Guzzo, T. (2015). Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications*, 125(3).
- Li, N., & Wu, D. D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision support systems*, 48(2), 354-368.
- Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining text data* (pp. 415-463). Springer, Boston, MA.
- Liu, B. (2010). Sentiment analysis: A multi-faceted problem. *IEEE Intelligent Systems*, 25(3), 76-80.
- Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *Ai & Society*, 30(1), 89-116.
- Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, 5(1), 80-92.
- Taboada, M., Brooke, J., Tofloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307.
- Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM computing surveys (CSUR)*, 34(1), 1-47.
- Ren, R., Wu, D. D., & Liu, T. (2018). Forecasting stock market movement direction using sentiment analysis and support vector machine. *IEEE Systems Journal*, (99), 1-11.
- Hoogendoorn, M., Berger, T., Schulz, A., Stolz, T., & Szolovits, P. (2017). Predicting social anxiety treatment outcome based on therapeutic email conversations. *IEEE journal of biomedical and health informatics*, 21(5), 1449-1459.
- Tai, C. H., Tan, Z. H., & Chang, Y. S. (2016). Systematical approach for detecting the intention and intensity of feelings on social network. *IEEE journal of biomedical and health informatics*, 20(4), 987-995.
- Scharl, A., Herring, D., Rafelsberger, W., Hubmann-Haidvogel, A., Kamolov, R., Fischl, D., ... & Weichselbraun, A. (2017). Semantic systems and visual tools to support environmental communication. *IEEE Systems Journal*, 11(2), 762-771.
- Sindhu, C., Vyas, D. V., & Pradyoth, K. (2017, April). Sentiment analysis based product rating using textual reviews. In *2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)* (Vol. 2, pp. 727-731). IEEE.
- Dhanalakshmi, V., Bino, D., & Saravanan, A. M. (2016, March). Opinion mining from student feedback data using supervised learning algorithms. In *2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC)* (pp. 1-5). IEEE.
- Guellil, I., & Boukhalifa, K. (2015, April). Social big data mining: A survey focused on opinion mining and sentiments analysis. In *2015 12th International Symposium on Programming and Systems (ISPS)* (pp. 1-10). IEEE.
- Claypo, N., & Jaiyen, S. (2015, January). Opinion mining for thai restaurant reviews using K-means clustering and MRF feature selection. In *2015 7th International Conference on Knowledge and Smart Technology (KST)* (pp. 105-108). IEEE.
- Schölkopf, B., Smola, A., & Müller, K. R. (1997, October). Kernel principal component analysis. In *International conference on artificial neural networks* (pp. 583-588). Springer, Berlin, Heidelberg.

AUTHORS PROFILE



Harjeet Kaur is an M.tech Student at Department Of C.S.E, Sachdeva Engineering College for Girls,Gharuan(Mohali),Punjab,India. She has received her B.Tech degree in Computer Science And Engineering from Adesh Institute Of Technology,Gharuan(Mohali) in the year of 2016 and she is pursuing her M.tech from SECG,gharuan presently.She is doing her research in the field of Data mining and Machine learning,titled as "Dimensionality Reduction in Sentiment Analysis using Colony-Support Vector Machine".Beside this ,her other research interests include network security,information security,web-designing and development and data analytics.She has research publication in UGC also.



Er. Prabhjeet Kaur is an Assistant Professor in Dept. of Computer Science And Engineering, at Sachdeva Engineering College for Girls, Gharuan. She has completed her B.Tech from Rayat and Bahra Engineering College(IKGPTU),Kharar(Mohali),Punjab,India and M.Tech from Punjabi University,Patiala. Her Specialization is in Machine Learning and Data Mining. She has research publications in reputed journals.