

¹Mohammad
Irsad*²Ashish Khare

Sentiment-Bearing String Extraction Method for Sentiment Analysis



Abstract: - Consumers are increasingly being asked for comments, feedback or suggestions regarding the products they purchase by e-commerce platforms. This recent trend has led to a noticeable increase in the amount of online product reviews along with the significant growth in online shopping. Other potential consumers are strongly influenced by the recommendations or critiques shared by other fellow consumers while making purchase decisions. Sentiment analysis (SA) records every user's feelings, views, beliefs and opinion regarding the specified product in order to determine if the general mindset of the user is positive, negative or neutral. It enables potential customers to take purchase decisions and enterprises to make the necessary improvements to the product for better satisfaction of their consumers. In this paper, a method to extract sentiment bearing strings has been proposed for sentiment classification on Amazon Mobile review dataset which extracts the sentiment bearing strings from the review and classifies the review based on computed sentiment score. It applies four machine learning classifiers namely Random Forest (RM), Logistic Regression (LR), KNN and SVM. The experimental results show that by applying proposed method the classifiers achieve much improved classification performance in terms of precision, recall and F1-score and accuracy. The SVM achieves highest accuracy of 97.69%. The proposed method performed better in the experimental evaluation than the current approaches when it compared to others.

Keywords: Sentiment Analysis, Natural Language Processing (NLP), Machine Learning

1. INTRODUCTION

The advancement in internet and web technology has affected nearly every aspect of human life. Social media and review sites have made the user very active and also provided a platform where people can freely share their opinions, views, and experience in the form of comments, reviews and feedback which can be valuable for other users. A huge amount of users generated contents are being posted daily on web. Organizational or individual decision-making is being impacted by the massive influx of information. Therefore it became very necessary to analyze user's view. The process to analyze the users views, opinions, aspects, experiences, feeling etc. with respect to objects such as products, issues, services etc and their related properties based the users comments, feedback and suggestions is called as sentiment analysis. It is a natural language processing (NLP) technique that uses computational methods to extract, classify, and quantify emotions and opinions expressed in text.

Sentiment analysis is also termed as opinion mining, review mining, sentiment mining, opinion extraction, affect analysis, subjectivity analysis, emotion analysis etc. [1]. The objective of sentiment analysis is to identify the sentiment behind written text and to categorize it as positive, negative, or neutral by applying classification algorithms [11].

Sentiment analysis has been categorized into following three categories:

1.1 DOCUMENT LEVEL

In document level sentiment analysis, the entire document is considered to express the sentiment of single object or entity [8]. Document level sentiment analysis assigns positive, negative or neutral opinion to the whole document and thus classifies the document as positive, negative or neutral for an attribute, product or service. It is not suitable for the document having comparison of multiple attributes [1].

1.2 SENTENCE LEVEL

The sentiment analysis at sentence level expresses sentiment of the sentence. The text paragraph is segmented into sentences and then polarity of the each sentence is represented as positive, negative or neutral for a specified product or service [3,5,6]. The sentence is generally classified into subjective and objective sentence.

¹ *Centre of Computer Education and Training, Institute of Professional Studies, University of Allahabad, Prayagraj, 211002 (India), Email: irshad_khan209@yahoo.com

²Department of Electronics & Communication, JK Institute of Applied Physics and Technology, University of Allahabad, Prayagraj 211002 (India), Email: ashishkhare@hotmail.com

Sentence level sentiment analysis is associated with subjective sentences and performs subjective classification which differentiates subjective sentences and objective sentences. The subjective sentence is a sentence which refers the opinion or feeling of a user for an object such as product, service or entity. Objective sentences are the sentences which express the factual information about an object such as historical data or facts. The objective assessment of a user on an issue is based on verifiable facts, statistics and other incontrovertible proof and does not take user's own feelings into account.

1.3 ASPECT LEVEL:

Aspect level sentiment analysis is a type of sentiment analysis in which the opinion described for all aspects of the entity are extracted. Aspect level analysis task is more detailed and complex than document level and sentence level sentiment analysis tasks. It is a finer-grained sentiment analysis task that analyzes and predicts the sentiment polarity associated with each aspect described in the text.

Apart from above three levels of sentiment analysis, an opinion is classified into two categories- regular opinion and comparative opinion. In regular opinion, the text directly expresses the sentiment on an object or on particular aspect of the object, for e.g. "voice quality of Samsung phone is amazing". In this example positive opinion is expressed for the aspect voice quality of object Samsung phone. In comparative opinion, the opinion is expressed through making a comparison between related entities or in terms of comparing multiple related aspects of two or more entities [2]. For e.g. "The voice quality of Samsung is better than that of Moto X.". This text expresses a comparison between Samsung and Moto X on the basis of voice quality and does not represent positive or negative opinion for voice quality but usually represents a relative ranking of two smart phones in terms of voice quality.

The application of sentiment analysis is domain dependent. Sentiment classification is very sensitive to the domain from which the training data is collected. Sentiment of a word used in one domain need not have same sentiment in other domain, apart from the words the format of the language for opinion expression also differs from one area of application to another. Therefore, if a classifier is trained using dataset from one field typically performs differently when tested or used to dataset from another area. So it becomes confusing when same word expresses positive opinion in one context but negative opinion in another context. As a result modification in domain is required. It is observed that earlier researches have employed tagged data from source domain, untagged data from the target domain and generic sentiment words as modification feature [4,7,9].

Presently, almost business groups or organizations of every sector are conducting some form of digital transformation which results a huge amount of unstructured and structured data. The major challenge for businesses is to convert large amount of unstructured data into useful insights that may help them to make massive decisions [10]. The results of sentiment analysis method are used effectively in customer services, financial systems, market research, reputation management, recommender systems, social media monitoring etc. Sentiment analysis can provide important insights by analyzing a huge amount of data and enables organizations to make data-driven decisions and improve their products and services. The impact of sentiment analysis result is not only limited to the field of natural language processing(NLP), but it also covers all other field of studies which include public opinion as main factor for decision making such as social science, political science, public relations, management studies etc.

The main contributions of the proposed work are as follows:

- It extracts the strings having sentiment from the review text and assigns a numeric value to each string having sentiment.
- It classifies the review either positive or negative on the basis of the overall calculated sentiment score of the review

The rest of the paper is organized as follows. Section 2 describes the important research papers related to the study of sentiment analysis. Section 3 describes dataset and the proposed methodology. The result and discussion is explained in Section 4. In the last, the conclusion is given in Section 5.

2. RELATED WORK

Product reviews are also considered as online reviews that are freely posted by the consumers on e-commerce sites which express their satisfaction level. Product reviews are the feedback provided by the consumers after product use. These reviews are important source of information for firms seeking to consumers and their requirements [12]. Reviews and feedback provide an engagement between consumer and business organizations which helps businesses to establish a relationship with the consumers. Product reviews have an impact on customer involvement as well [22]. Hu, M. and Liu, B, in 2004, proposed a set of mining techniques to summarize product reviews based on natural language processing and data mining techniques [13], considering the goal to give an abstract of larger number of consumer reviews of a certain online business. It was their assumption that the review summarization is important as more consumers buy products online and express their feedback, feeling, or views about the product in the form of reviews on the Internet. The ability to summarize reviews is beneficial not only to consumers, but also to manufacturers and venders of the product. Vanaja S, BelwalM.[14] proposed aspect-level sentiment analysis using identification, clustering and classification. They have used preprocessing techniques to remove stop words and other unwanted words applied POS tagging to extract adjectives from the sentences. For classification, NB and SVM classifiers have been used. They suggested that SVM classifier was not suitable for large datasets. Jagdale RS et al. [15] have performed sentiment analysis on a particular product review using NB and SVM approaches. They applied preprocessing technique to clean data and used Bag-of-Words model for feature extraction. The dataset was compared to opinion lexicons, which included 4783 negative and 2006 positive terms with opinion score assigned to each sentence. Machine learning classifiers NB and SVM were used with opinion score and various features, and the accuracy was calculated. For mobile phone review, NB achieved the accuracy of 92.85% and SVM also achieved 92.85% accuracy.

Barkha Bansal and Sangeet Srivastava [16], proposed a Hybridized Attribute-centric Sentiment Classification (HABSC) to incorporate the knowledge of a specified domain and to gather connections of implicit words. This method identified the most common bi-grams and tri-grams in the dataset, followed by POS tagging to preserve sentiment words and aspect characteristics. In addition, this approach performed topic extraction and used TFIDF to represent each document. By using pre-defined lexicon and knowledge of a specified domain, all the adjectives and adverbs were tagged. The result demonstrated that the accuracy of the classification of HABSC model is much better as compared to other approaches. Additionally, it takes less time to compute as compared to distributed vectorization framework. It is somewhat complex to build, improve and implement a functional method of natural language processing (NLP). As a result, Convolution Neural Network (CNN) is being frequently adopted for NLP by the researchers [17]. This offers better results than machine learning models. Most of the present researches have used dependency tree to concentrate on learning dependency from contextual words to aspect word. It ignores contextual emotional knowledge with reference to the particular aspect to take advantage of the emotional dependences of the sentence based on the particular aspect, graph convolutional network based on SentiNet has been proposed by Liang et al.[18].

N. Nandal et. al. [22] have presented a novel approach sentiment detection at aspect level by focusing the features of products. They utilized crawled Amazon customer reviews for their study. They identified the aspect terms present in each review. To extract appropriate information from the review dataset, the their system performed different pre-processing steps like tokenization, stemming, casing and removal of stop words. At last, the system assigns a rank based on whether the information is classified as positive or negative.

3. PROPOSED METHODOLOGY

This section describes the methodologies used in this work for sentiment analysis. The flow of the steps for proposed work are described in Figure-1

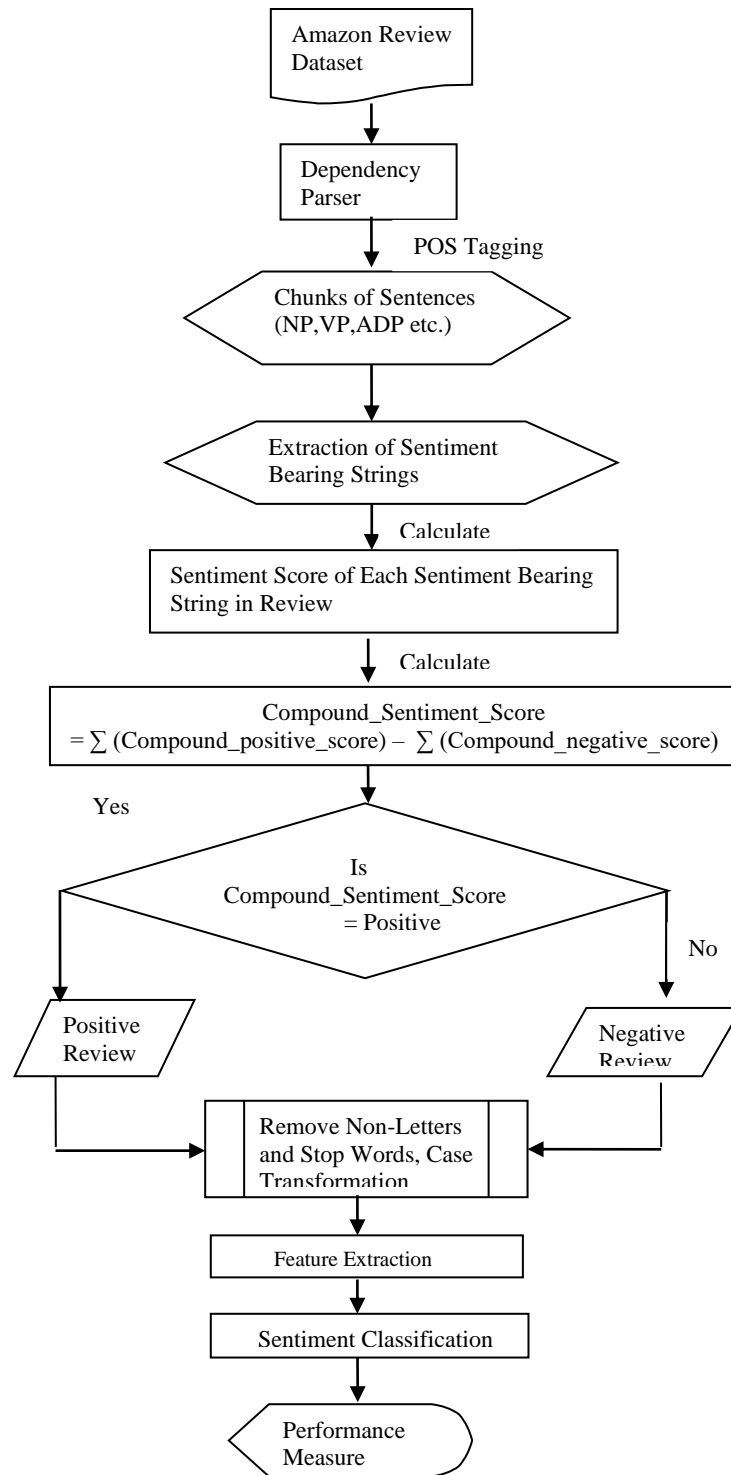


Fig.1. Flow of Steps of Proposed Methodology

3.1 Dataset description

In this research work, Amazon Mobile review dataset has been used. The dataset contains the following properties as columns.

asin: This column represents product ASIN (Amazon Standard Identification Number) which is unique for each mobile or mobile variable.

name: This column represents the name of the reviewer.

rating:This column represents the rating of the reviewer for mobile on scale of 1 to 5.

date: date of the review.

verified:This column represents valid customer as true or false. The dataset contains the reviews from 90% valid users.

title:Represents tile of the review.

body:This column represents the review contents of the reviewer.

helpfulness:helpfulness feedback.

The dataset, Amazon Mobile Review has been downloaded from Kaggle. There are 67987 data samples in dataset. In the first step, important features are extracted from dataset and the features which have not to be used are removed. Finally, the dataset has only two columns: body and rating.

The rating column contains the rating in numeric form given by the consumer corresponding to each mobile review and it has the range from 1 to 5. On the basis of the rating values, we have calculated sentiment score and added one more column named as 'Sentiments'. The sentiment has been calculated as-

- If rating value ≥ 3 then sentiment is positive and it is labeled as 1 in the Sentiments column.
- If the rating value < 3 then sentiment is negative and it is labeled as 0 in the Sentiments column.

As a result, we have tagged dataset for the task of sentiment analysis. To clean each review, we applied preprocessing techniques on each mobile review.

3.2 Method

The dataset is downloaded from Kaggle and used in this proposed work. The dataset is first passed in to dependency parser. The methodology involve following steps:

a. Dependency Parsing:

It is a Natural Language Processing (NLP) technique that recognizes word relationships by examining the grammatical structure of a sentence. It entails determining how one word in a sentence relates to another, thereby helping in comprehending the arrangement and significance of the text. Dependency parsing is a helpful tool in process of sentiment analysis since it assists in determining whether words modify or are connected to sentiment-bearing words by analyzing structure of the sentence. It makes simpler to determine a text's general sentiment more precisely.

b. Part-of-Speech (POS) tagging

Part-of-Speech (POS) tagging involves categorizing each word in the text according to its suitable grammatical category. Words are divided into groups including verbs, adjectives, adverbs, and nouns. Determining the significance of each word in the phrase and proceeding through the sentiment analysis process depends on this stage of the process. Based on the POS tags, the words are grouped into chunks. The common types of chunks include:

Noun Phrase (NP): It is a noun-functioning phrase, usually made up of a noun and its modifiers.

Verb Phrase (VP): It is a verb-functioning phrase that usually consists of the verb along with any complements, modifiers, and auxiliaries.

Adposition Phrase (ADP): It is a prepositional or postpositional phrase that frequently includes both an adposition and its complement.

Chunking helps in identifying the main components of a sentence and the relationships between components.

c. Extraction of Sentiment Bearing Strings

Identifying textual sequences that express sentiment is performed in this step. Phrases or words that convey a positive, negative, or neutral feeling are known as sentiment-bearing strings. These may be single terms like "wonderful" or "amazing," or they could be sentences like "not impressed" or "strongly suggested."

d. Sentiment Score calculation:

Sentiment-bearing strings are assigned sentiment scores that indicate the emotional tone of the text. These scores can be positive, negative, or neutral, reflecting the overall sentiment expressed. Typically, a numerical value represents the intensity of the sentiment, the higher positive sentiment is represented by higher positive sentiment values of sentiment whereas lower negative values indicating strong negative sentiment, and the sentiment scores close to zero representing neutral sentiments. To achieve this goal, VADER (Valence Aware Dictionary and sEntiment Reasoner) is used. Vader is a rule-based lexicon and sentiment analysis tool which intended to identify sentiments expressed in social media, although it also performs well in product reviews and ordinary text. In addition to the positivity, neutrality, and negativity ratings, VADER also generates a compound score, which is a single metric that expresses the overall attitude determined by combining the preceding three. The highest and lowest values of these scores are -1 (very negative) and +1 (very positive). For each string in the review that conveys sentiment, the compound positive sentiment scores and compound negative sentiment scores are added together respectively and finally the negative sentiment scores are subtracted from the positive sentiment score to get the final compound sentiment score. This provides a review's total sentiment score which is calculated as following;

$$\text{Compound Sentiment Score} = \text{Sum of Compound Positive} - \text{Sum of Compound Negative} \dots(1)$$

On the basis of compound sentiment score, a review is categorized into positive or negative. This process has been performed for all the reviews of dataset to calculate the compound sentiment score and based on this score, a class is given to the review. If the compound sentiment score of a review is positive, then the review is classified as positive. Otherwise, it is classified as negative.

3.3 DATA PREPROCESSING

Data preprocessing is an important step in the process of sentiment analysis. The goal of preprocessing is to make data ready for analysis and prepare the data in such a way that it can be easily understood and analyzed to extract valuable insights. It involves following steps-

3.3.1 Data cleaning and tokenization

The review data is collected from online sources which generally contain irrelevant information such as HTML tags, URLs, numbers, scripts and special characters. In data cleaning process, the irrelevant data is removed and only text is kept. This process ensures that the algorithm focuses on valuable information only. The text document is a collection of words, every individual word is called a token. Tokenization is a process of splitting the text into tokens.

3.3.2 Case Transformation

Review text contains the terms in lower case and upper case both. It is required to convert all the terms into same case. This process standardizes the review text by converting all characters of review text into lower case.

3.3.3 Stemming

It is the process to convert a word into its root or base form. It involves the process of clipping the affixes from the word to make it shorter as simply possible, yet different words refer to the same meaning after stemming.

3.3.4 Data filtering

The stop words do not convey any sentiment so these words are required to be removed. The filter function eliminates English stop words from a text by eliminating each word that matches a word from a built-in list of English stop words.

3.4 FEATURE EXTRACTION

The dataset contains a large amount of data sometimes data is redundant and too vast to process in a feasible amount of time with feasible resources. So it may be complex to extract suitable information from data. Data representations a major task to represent data in an arranged way for a machine learning algorithm to run in a feasible amount of time. For this goal to achieve the dimension of the data set is reduced into manageable groups for further processing. This is called feature extraction. By applying the feature extraction process, the amount of duplicate data can be reduced and the extracted data are supposed to have suitable information, which can be used to perform the desired task with the reduced and managed data.

It is a major and necessary task to represent data in a defined format to build a model for machine learning algorithm. Bag of words (BoW) is a general feature representation method which is mostly used by machine learning algorithms. In a feature representation method, the features determine some specific properties that are considered to describe the underlying pattern of data. Before passing into a machine learning algorithm, each review of the entire dataset is required to be represented in terms of specified features. There are a large number of features like BoW, unigram, bigram, trigram, parts of speech (POS) etc. feature are frequently used for sentiment analysis task. A review text is represented in BoW as a collection of words and phrases present in the review, ignoring the grammar and order of the words but the frequency of the word is maintained. The frequency of each word is used as a feature for training the classifier.

Bag of Words (BoW) is a feature extraction method from text data. The review data is unstructured and a machine learning algorithm requires fixed length and structured input. By using BoW technique, the review text is converted into vectors of fixed length. In machine learning algorithms, selection of proper attributes is the first step to represent the instances of data. The training and testing data will be defined based on the selected attributes and thus attributes are represented by features. The dataset is transformed into different sets of unigram, bigram or trigram or their combinations for the process of attribute selection. A unigram is a phrase having sequence of one word. BoW is considered as equivalent to unigram. The choice has to be made for feature representation by using frequency and presence of the unigram, bigram, trigram or combination of unigram and bigram presence and frequency.

The feature vector must be chosen before training and testing of the classifier. A feature vector is a numerical representation of the distinctive characteristics of an object. Thus creation of an effective feature vector is crucial to the implementation of a classifier. For the proposed work, the feature vectors have been used like bigram and trigram. The feature vectors employed for the proposed work include characteristics like the existence of bigrams, trigram or a mix of both. Each data instance's feature words will combine to form the whole feature vector (review text). These feature vectors were used to train the classifier.

In order to test, finding the feature words that produce a different pattern of feature vector is required for evaluating a review text. This feature vector serves as the model's input for learning, and the classifiers use the learning to forecast or anticipate the emotion of the reviews.

4. RESULT AND ANALYSIS

In this section we have described the result of machine learning classifiers. By creating training and testing sets, supervised machine learning was utilized to develop sentiment classification models such as Random Forest (RF), Logistic Regression (LR), KNN and Support Vector Machine (SVM). Next, a set of features was extracted from the testing and training data and utilized in a classifier model. The entire dataset was divided into training and testing sets. To train the model, 80% data was used while rest of 20% data was used for evaluating the performance of the model.

The result indicates that the SVM achieves highest accuracy score of 89.92%. We have used sentiment bearing string extraction approach and then applied this approach with machine learning classifiers for review classification. After applying sentiment bearing string extraction approach, we found that the results of classifiers are better as compared with the traditional approaches. The classification result indicates that the SVM achieves highest accuracy score of 97.69% and LR achieves 96.83 accuracy score. These classifiers perform best for the large feature set. The accuracy of RF classifier is also better with the accuracy score of

96.21%. The performance of KNN is poor with accuracy score of 89.58%. The experimental result is given in Table 1 and Table 2.

Table.1. Classification result of machine learning classifiers

Classifier	Class	Precision	Recall	F1-Score	Accuracy (%)
Random Forest (RF)	Negative	0.83	0.71	0.77	89.21
	Positive	0.91	0.95	0.93	
Logistic Regression (LR)	Negative	0.83	0.74	0.78	89.77
	Positive	0.92	0.95	0.93	
KNN	Negative	0.73	0.11	0.20	77.08
	Positive	0.77	0.99	0.87	
SVM (linear)	Negative	0.82	0.76	0.79	89.92
	Positive	0.92	0.94	0.93	

Table.2. Classification result of machine learning classifiers using sentiment bearing string extraction approach

Classifier	Class	Precision	Recall	F1-Score	Accuracy (%)
Random Forest (RF)	Negative	0.92	0.85	0.88	96.32
	Positive	0.97	0.99	0.98	
Logistic Regression (LR)	Negative	0.96	0.84	0.90	96.83
	Positive	0.97	0.99	0.98	
KNN	Negative	0.68	0.84	0.73	89.58
	Positive	0.97	0.91	0.94	
SVM (linear)	Negative	0.95	0.91	0.93	97.69
	Positive	0.98	0.99	0.99	

The experimental result shows that RF, LR and SVM classifiers provides better performance as compared to KNN in both the cases. After applying the sentiment bearing word extraction approach, the performance of RF and LR is improved by 7%, SVM by 8% and KNN by 12%. The improved performance is attributed due to sentiment bearing string extraction approach. The irrelevant strings which do not bear any sentiment are removed and therefore review contains only the strings which bear sentiment either positive or negative. A significant improvement in the performance of KNN is also observed. The accuracy of the KNN is increased from 77.08% to 89.58%. Figure 2 shows the classification performance of the classifiers in terms of accuracy.

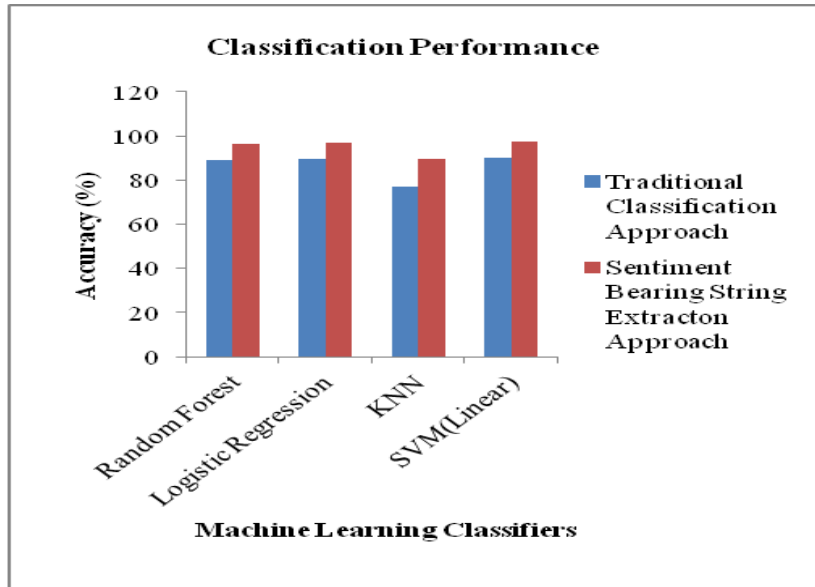


Fig.2. Performance of Machine Learning Classifiers

The review is classified into positive and negative categories. The proposed approach is applied on both the classes of review. Figure 3 show the performance measures of negative reviews in terms of precision, recall and F1-Score of machine learning classifiers before applying the proposed approach. After applying the proposed approach, a significant improvement has been observed in the values of these performance measures which is shown in Figure 4.

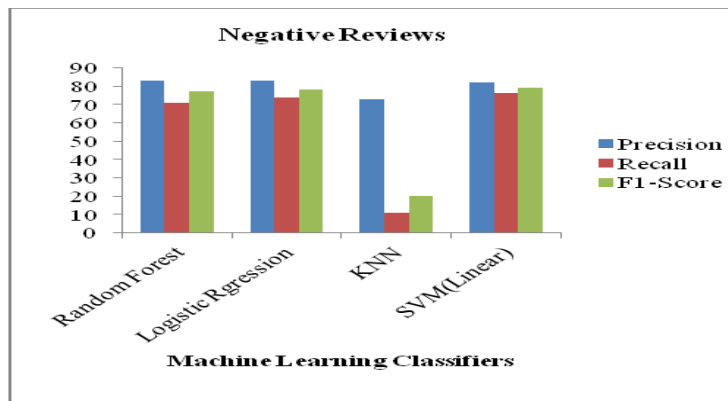


Fig.3. Precision, Recall and F1-Score of Machine Learning Classifiers on negative reviews

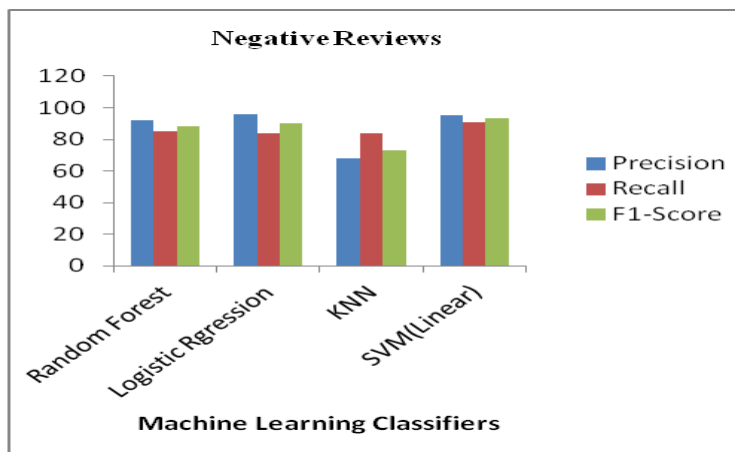


Fig.4. Precision, Recall and F1-Score of Machine Learning Classifiers on negative reviews after applying sentiment bearing string extraction approach

Figure 5 and Figure 6 show the value of precision recall and F1-Score of classifiers of positive reviews before and after applying the proposed approach.

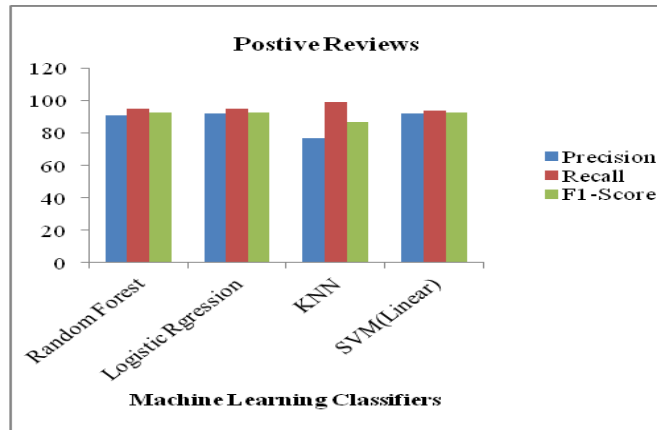


Fig 5. Precision, Recall and F1-Score of Machine Learning Classifiers on positive reviews

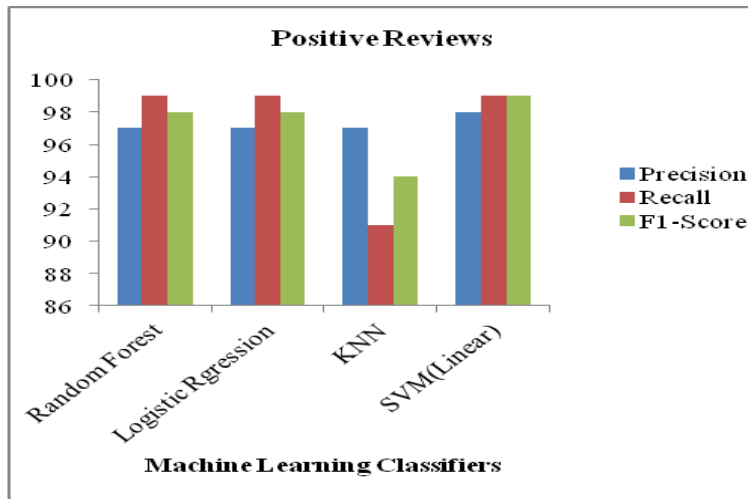


Fig. 6. Precision, Recall and F1-Score of Machine Learning Classifiers on positive reviews after applying sentiment bearing string extraction approach

We have also compared our result with previous researches and observed that the performance of proposed work is better. In paper [15], two classifiers namely Naïve Bayes (NB) and SVM has been applied on Amazon Mobile Phone Review dataset and both the classifiers achieve the accuracy of 92.85%. In paper [19], four different classifiers are used and then fusion approach has been developed by combining SVM and Decision Tree classifiers. They achieved the highest classification accuracy of 93.4 by using fusion approach. In paper [20], traditional machine learning classifiers named as Logistic Regression, Naïve Bayes and SVM along with deep learning model named as RNN have been used for classification task and they achieved the highest frequency of 91%. Deep learning model CNN along with word2vec feature extraction method has been used in paper [21] which achieves the classification accuracy of 88%.

By comparing the result of the proposed work with previous works, it is observed that the performance of the machine learning classifiers are better along with the proposed approach. Table 3 shows the comparison between proposed work and previous works.

Table.3. Comparison of proposed work with previous researches

Reference	Dataset	Classifier/Model	Accuracy (%)
[15]	Amazon	NB	92.85

	Mobile Phones Review	SVM	92.85
[19]	Mobile Review	Decision tree	88.6
		SVM	89.3
		Naïve Bayes	86
		Random Forest	87.9
		Fusion approach	93.4
[20]	Amazon Smartphone Reviews	Logistic Regression	89
		Naïve Bayes	87
		SVM	89
		RNN	91
[21]	Analysis on Mobile Phone Brands Reviews	Word2vec-CNN	88
Proposed work	Amazon Mobile Reviews	Random Forest (RF)	96.32
		Logistic Regression (LR)	96.83
		KNN	89.58
		SVM (linear)	97.69

5. CONCLUSION

In the proposed work, an approach to extract sentiment bearing strings form the review for sentiment analysis has been proposed. This paper implemented four baseline machine learning classifiers- Logistic Regression, Random Forest, KNN and SVM and evaluated the performance of the classifiers in terms of precision, recall, F1-score and accuracy. The performance of the classifiers are evaluated again after implementing the proposed approach and it is observed that performance of all the classifiers are better. The accuracy measurement of the classifiers is given in Table 2. SVM achieved the highest accuracy of 97.69%. This work can be extended for aspect based sentiment analysis in future to extract the aspects and their sentiments. By using aspect based sentiment analysis, we can get the sentiment of the every aspect represented in the reviews.

REFERENCES

- [1] Liu, B.: Sentiment Analysis and Opinion Mining, p. 7. Morgan and Claypool Publishers, USA (2012)
- [2] Jindal, N., Liu, B.: Mining comparative sentences and relations. In AAAI22, 1331–1336 (2006)
- [3] Meena A, Prabhakar TV (2007) Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis. In: Amati G, Carpineto C, Romano G (eds) Advances in information retrieval. ECIR 2007. Lecture notes in computer science, vol 4425. Springer, Berlin, Heidelberg.
- [4] Aue, A., Gamon, M.: Customizing Sentiment Classifiers to New Domains: A Case Study. In: Proceed-ings of Recent Advances in Natural Language Processing (RANLP-2005) (2005)
- [5] Arulmurugan R, Sabarmathi K, Anandakumar H (2019) Classification of sentence level sentiment analysis using cloud machine learning techniques. Cluster Comput 22(1):1199–1209
- [6] Shirsat VS, Jagdale RS, Deshmukh SN (2019) Sentence level sentiment identification and calculation from news articles using machine learning techniques. In: Computing, communication and signal processing. Springer, pp 371–376
- [7] Blitzer, J., Dredze, M., Pereira, F.: Biographies, Bollywood, Boomboxes and Blenders: Domain Adaptation for Sentiment Classification. In: Proceedings of Annual Meeting of the Association for Computational Linguistics (ACL-2007) (2007)

- [8] Pu X, Wu G, Yuan C (2019) Exploring overall opinions for document level sentiment classification with structural SVM. *Multim Syst*25(1):21–33
- [9] Pan, S., Ni, X., Sun, J., Yang, Q., Chen, Z.: Cross-domain Sentiment Classification via Spectral FeatureAlignment. In: *Proceedings of International Conference on World Wide Web (WWW-2010)* (2010)
- [10] Ahmad Z, Jindal R, Ekbal A, Bhattacharyya P (2020) Borrow from rich cousin: transfer learning for emotion detection using cross lingual embedding. *Expert Syst Appl* 139:112851
- [11] Priyadarshini I, Cotton C (2021) A novel LSTM-CNN-grid search-based deep neural network for sentiment analysis. *J Supercomput* 77(12):13911–13932.
- [12] Zhou F, Jiao R J, Linsey J S. Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews [J]. *Journal of Mechanical Design*, 2015, 137(7): 071401-1-071401-12.
- [13] Hu, M., Liu, B.: Mining and Summarizing Customer Reviews. In: *Proceedings of the tenth ACM SIGKDD international conference on Knowledge Discovery and Data Mining* (168–177). ACM (2004).
- [14] Vanaja S, Belwal M. Aspect-level sentiment analysis on e-commerce data. In: *international conference on inventive research in computing applications (ICIRCA)*, IEEE, pp. 1275–1279, 2018.
- [15] Jagdale RS, Shirsat VS, Deshmukh SN. Sentiment analysis on product reviews using machine learning techniques. In *Cognitive Informatics and Soft Computing*, Springer, Singapore, pp. 639–647, 2019.
- [16] Bansal Barkha, Srivastava Sangeet. Hybrid attribute based sentiment classification of online reviews for consumer intelligence. *Appl Intell.* 2019;49(1):137–49.
- [17] Lopez, M. M., Kalita, J. Deep Learning applied to NLP. arXiv:1703.03091. 2017 Mar 9.
- [18] Liang, B., Hang, S., Gui, L., Cambria, E. & Ruifeng, X. Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowl.-Based Syst.* 235, 107643 (2022).
- [19] Chhaya Chauhan, Smriti Sehgal. "Sentiment Classification for Mobile Reviews using KNIME". 2018 International Conference on Computing, Power and Communication Technologies (GUCON) Galgotias University, Greater Noida, UP, India. Sep 28-29, 2018
- [20] Neelesh Sharm, Tarun Jain, Saket S Narayan, Anurag C Kandakar "Sentiment Analysis of Amazon Smartphone Reviews Using Machine Learning & Deep Learning". In 2022 IEEE International Conference on Data Science and Information System (ICDSIS) 2022.
- [21] Noor Izati Abdul Hamid, N. Kamal, H. M. Hanum, Noor Latiffah Adam, Z. Ibrahim. "fProSentiment Analysis on Mobile Phone Brands Reviews using Convolutional Neural Network (CNN)". *Computer Science 2022 IEEE International Conference on Computing (ICOCO) 2022*
- [22] N. Nandal, R. Tanwar, and J. Pruthi, "Machine learning based aspect level sentiment analysis for Amazon products," *Spatial Information Research*, vol. 28, no.5, pp.601–607, Oct.2020, <https://doi.org/10.1007/s41324-020-00320-2>.