iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48678

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Sentiments Analysis of Amazon Reviews Dataset By using Machine Learning

Varsha Gupta[1], Prof. Vijay Prakash Sharma[2]
[1]*M.Tech Research Scholar,* [2]*Assistant Professor, Shri Ram College of Engineering, Banmore, India*

Abstract: ***Any opinion of a person that can convey emotions, attitudes, or opinions is known as a sentiment. The data analyzes that are collected from media reports, consumer ratings, social network posts, or microblogging sites are classified as opinion mining research. Analysis of sentiment should be viewed as a way of evaluating people for particular incidents, labels, goods, or businesses. The amount of views exchanged by people in micro-logging sites often increases, which makes nostalgic interpretations more and more common today. All sentiments may be categorized as optimistic, negative, or neutral under three groups. The characteristics are derived from the document term matrix using a bi-gram modeling technique. The sentiments are categorized among positive and negative sentiments. In this analysis, the Python language is used to apply the classification algo for the data obtained. The detailed accomplishment of LinSVC demonstrates greater precision than other algos.***
Keywords: Sentiment Analysis, Classification Techniques, Naïve Bayes, Voting, Linear SVC , NLP,Data mining, amazon product review dataset.

I. INTRODUCTION

The current Internet era has been an enormous cyber database that houses vast amounts of data that users generate or use. The database has expanded at an unprecedented rate, generating a digital market of consumers sharing their opinions on Facebook, Twitter, Rotten Tomatoes, and Foursquare. Opinions shared as comments offer new study tools to recognize the mutual desires or dislikes of cyber societies. The category of analysis that impacts everyone from the viewer, film reviewers to the production team, is one of these areas of reviews. The film reviews on the blogs are not systematic reports, but rather casual and unstructured. The viewpoints conveyed in film reviews represent quite accurately the sentiment transmitted. The inclusion of such broad usage of words to describe the revisions inspired us to evaluate the polarity of the film in such terms of feeling. Sentiment Analysis is a technology that will be relevant over the next few years. With opinion processing, we can differentiate bad content from high-quality material. Through current technology, we will learn if a film has better views or poor views and why these views are good or negative. In this area, a significant proportion of early work centered on user feedback, such as feedback on Amazon.com[1], describing sentiments as favorable, negative, or neutral. The majority of sentiment analysis studies currently rely on social networking sites like IMDB, Twitter [2] & Facebook, which need the correct methods to satisfy through text demand. In comparison, the study of the sentences in film reports is a difficult task. ResearchSentiment Mining is a process focused on the NLP or information extraction (IE) approaches to review a broad variety of documents such that the views of different writers can be collected[3]. This method includes a variety of techniques, including machine etymology and IR [4]. The fundamental principle of sentiment analysis is to recognize and define the polarity of text or short messages. True, "negative," or "impartial" (neutral) opinion polarity is classified.It must be emphasized that emotion mining may be carried out in the following three stages. [5]:

1) Classification of sentiment at a document level: atthis point, a text may completely be categorized as "positive" or "neutral."
2) Classification of sentences: Each sentence shall be categorized as 'yes,' 'no' or 'unbiased.'
3) Sensitivity classification of dimension or type of features: at this point, phrases/documents may be categorized as « positive », « negative » or « non- party » provided those aspects of sentences/archivesbut generally recognized as « view grouping of the viewpoint stage.

II. REVIEW OF LITERATURE

The paper research that we present takes account of previousstudies in the issue field of opinion analysis in film reviews. Stefano, Andrea & Fabrizio in [6]Present SentiWordNet 3.0, a lexical tool particularly built for the classification of sentiment. SentiWordNet 3.0 is a science collaborative platform and has been approved in over 300 study groups at present.Senti WordNet 3.0 focuses especially on aspect enhancements (the algos used for WordNet annotation) it includes. Godbole, Manjunath & Stevens in their work [7]Current to each distinct entity in text corpus a framework which quantifies positive or negative opinion.

©IJRASET: All Rights are Reserved

It consists of twostages, one stage in which sentiment perception is decided byperceptions and one step in which a relative score is defined for each person. Annett &Kondark [8]It was noted that the ML sentimental classification methodology on film review is very effective and that the kind of features picked has a significant effect on the classifier's accuracy. Since there is an upperlimit to the exactness point, as is apparent in a dictionary- based method. Pang & Lee work [9]is known in sentimental movie review research as a standard. The problem is not graded by subject matter, but by general emotion, e.g. whether a comment is good or bad. They claimed that classical learning methods in machines have stronger performance than baselines made. Though, 3 approaches of machine learning (Naive Bayes, maximum entropy classification &SVM do not have such strong results in the classification of emotions as in the conventional subject classification. They also derive and introduce productive methods for the quest for minimum cuts in graphs; this significantly favors the fulfillment of cross sentences semantic constraints, which offer an excellent means of combining conceptual knowledge at an interdenticle level for conventional word dictionaries. Singh et al. [10]Present the theoretical research on the SentiWordNet results assessment method for classifying film reviews and blog posts on documentary feelings. Similar to two common machine- learning methods they made improvements in semantic functions, ranking schemes, and thresholds of the SentiWordNet method. Comparative outcomes in film reviews and forum posts are demonstrated by normal precision, F- measure, or Entropy efficiency assessment measures. Chunxu Wu [11]Proposed a system for integration of context-dependent views semantic by WordNet. To order to assess opinion by way of semantic closeness tests, the approach suggested is used. This methodology depends on these steps to assess the subject of feedback where inadequate knowledge is accessible. Taboada et al. [12]Used a strategy of lexicon-dependent identification and interpretation of documents depending on their emotions. Strong or derogatory lexicons have been used to do this correctly. Furthermore, the simulator of semantic orientation (SO-CAL) focused on intensifiers and negations is introduced. This method has reached an accuracy of 76.37 percent in film reviews. Zagibalov et al. [13]Addressed the problem of classifying the feelings in consumers in Chinese written goods. Their method relied on unattended classification underwhich the vocabulary seed was born. Initially, there was only one (good) term identified as optimistic. The first seed was retrained iteratively to identify emotions. The criteria for opinion density is then used to measure the feels ratio for a paper. The tests demonstrated that after 20 iterations, the qualified classifier obtained an 87% FSR for dramatic polaritydetection. Tripathy et al. [14]Tried to distinguish feedback by polarity with supervised learning algos like Naïve Bayes, SVM, random forest, or linear discriminant analysis. Inachieving so, four phases were included in the solution suggested. The first step was to delete stopping words, mathematical characters, and individual characters. Second, text comments have been translated into a numerical matrix. Thirdly, vectors created for four different classifiers were usedas inputs. Various parameters were then measured to assess the efficiency of the proposed n method, e.g. precision, alert, f-measurement, or classification accuracy. The random forest classifier surpassed many classifiers for Polarity & IMDb datasets. Saleh et al. [15]To distinguish paper evaluations, the SVM was applied to 3 separate datasets. To determine the effect of the SVM in classification papers, many n-gram schemes have been used. The researchers used three weighing methods to produce vectors of function: Term Frequency, Conditional Occurrence, and Term Frequency Inverse Document Frequency (TFIDF).

III. PROPOSED METHODOLOGY

The algo works with probabilistic, and the algo works with different measures. The data collection will be evaluated before analysis so each sentence in the dataset is split into terms so each word's polarity will be evaluated. The term polarity may be viewed as positive or negative in a tabulated form as seen in figure 1. The algorithm is introduced to the dataset following this procedure. Word level research is the strongest as it is most likely to produce reliable information. The next step is to implement algo suggested as each term is measured in two groups (positive or negative) with both positive and negative probability. The higher probability classification is defined as corresponding sentence probability.

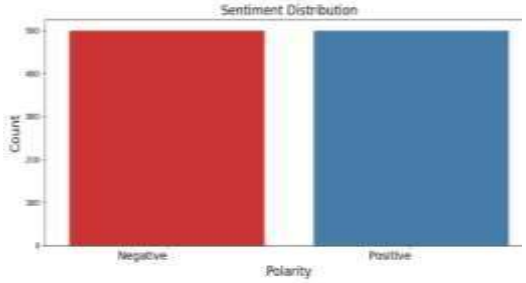


Figure 1. Visualization of the polarity of words as positive andnegative

A. Dataset Collection

This dataset contains user ratings and Amazon documentation as shown in figure 2. This data collection includes comments (ratings, text, votes for helpfulness), product metadata (descriptions, details in category, size, brand, or images) or links (also seen or bought graphs).

```
#reading dataset
df = pd.read_excel('amazon.xlsx')
print(df.head())
df.info()
polarity=df['Polarity'].values
```

```
                                              Review  Polarity
0  So there is no way for me to plug it in here i...         0
1                        Good case, Excellent value.         1
2                             Great for the jawbone.         1
3  Tied to charger for conversations lasting more...         0
4                                  The mic is great.         1
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 2 columns):
Review      1000 non-null object
Polarity    1000 non-null int64
dtypes: int64(1), object(1)
memory usage: 15.7+ KB
```

Figure 2. Representation of amazon dataset

B. Preprocessing

Preprocessing data is a technique in data mining that allows raw data to be converted into an understandable format. Real-world statistics are often incomplete, unreliable, or deficient in any action or patterns, and include several mistakes. Preprocessing data is a well-proven approach to overcome such problems. For more analysis, computer preprocessing prepares raw data. We removed unique characters and numeric values for our study and translated allletters into smaller instances. Snowball stemmer: a little string processing language for the development of stemming algos for knowledge retrieval. Snowball: Snowball Stemming effectively eliminates a term suffix or returns it to its source term. For eg, if we take "ing" from "flying" we get a term or a root term from "flying." "flying" is a verb, and the suffix is an "ing" "ing." This suffix is used to build a separate term fromthe initial. The word matrix of a concept is then defined asThe term frequency for each word is recorded in each case byconcept term. We begin with the Words Sack, which displays the documents and then count the number of occasions a word appears within each text. The next step is to split the data between preparation and assessments by splitting the data collection between 90% or 10%. The evaluation of data collection would be evaluated. The next step is to submit the classification algorithm and get the results. A vote is one of the easiest ways to merge several learning algos with predictions. Voting classification is not a classifier but a wrapper for different ones that have been equipped and rated concurrently to take advantage of the distinct characteristics of each algo.

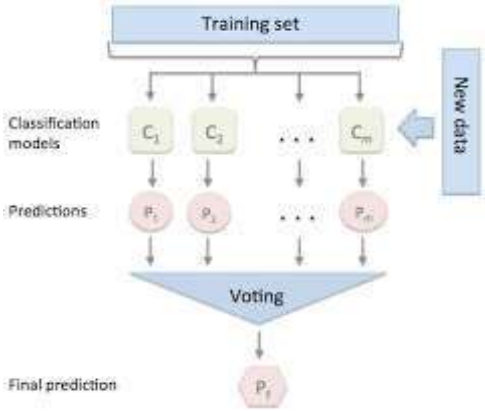


Figure 3 Voting Classifier

Linear SVC: A Linear SVC aims to match the data you have and return the "right fit" hyperplane, which separates or classifies the information. This is also the intention of a LinearSVC classification algo. You should then feed some elementsinto your classifier after getting hyperplane and see what "predicted" type is. This renders this specific algo more appropriate for our use.

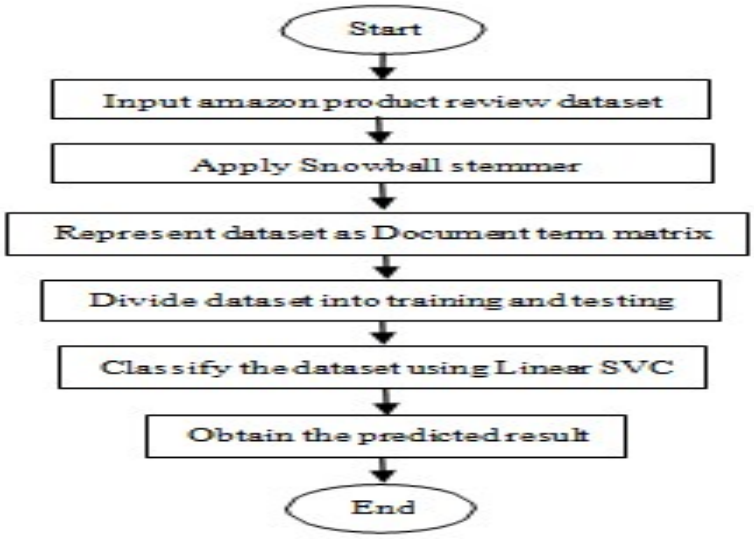


Figure 4 Flow diagram of the proposed methodology

IV. RESULTS AND DISCUSSION

Precision or recall levels are measured, and 91.66% or 89.79% of meaning are collected. The average accuracy is 91%. The precision of 91 percent indicates that LinSVC algo is higher than the other two algos in classification.

```
fpr, tpr, thresholds = metrics.roc_curve(y_test, predictions, pos
accuracy=accuracy_score(y_test, predictions)
precision=precision_score(y_test, predictions)
recall=recall_score(y_test, predictions)
f1 = f1_score(y_test, predictions)
auc=metrics.auc(fpr, tpr)
print("Naive bayes AUC: {:.4f}".format(auc))
print("Accuracy  : {:.4f}%".format(accuracy*100))
print("Precision : {:.4f}%".format(precision*100))
print("Recall    : {:.4f}%".format(recall*100))
print("F1-measure: {:.4f}%".format(f1*100))
```

```
Test Confusion Matrix :
[[42  7]
 [ 5 46]]
Naive bayes AUC: 0.8796
Accuracy  : 88.0000%
Precision : 86.7925%
Recall    : 90.1961%
F1-measure: 88.4615%
```

Figure 5 Result visualization of Naïve Bayes

AUC — ROC is a classification problem output calculation at various levels. ROC is a chance curve and AUC is a degree or separability metric. This asks how many models will discriminate between groups. As AUC is larger, the stronger the pattern is 0s as zero and 1s as 1s.

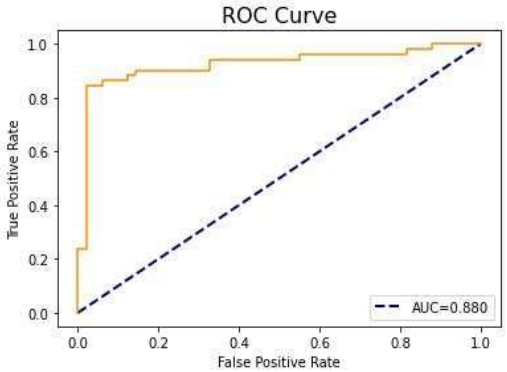


Figure 6 ROC curve of Naïve Bayes

```
fpr, tpr, thresholds = metrics.roc_curve(y_test, predictions, pos_label=1)
accuracy=accuracy_score(y_test, predictions)
precision=precision_score(y_test, predictions)
recall=recall_score(y_test, predictions)
f1 = f1_score(y_test, predictions)
auc=metrics.auc(fpr, tpr)
print("Voting AUC: {:.4f}".format(auc))
print("Accuracy  : {:.4f}%".format(accuracy*100))
print("Precision : {:.4f}%".format(precision*100))
print("Recall    : {:.4f}%".format(recall*100))
print("F1-measure: {:.4f}%".format(f1*100))
```

```
Test Confusion Matrix :
[[40  7]
 [ 3 41]]
Voting AUC: 0.9034
Accuracy  : 90.0000%
Precision : 85.4167%
Recall    : 93.1818%
F1-measure: 89.1304%
```

Figure 7 Result visualization of Voting

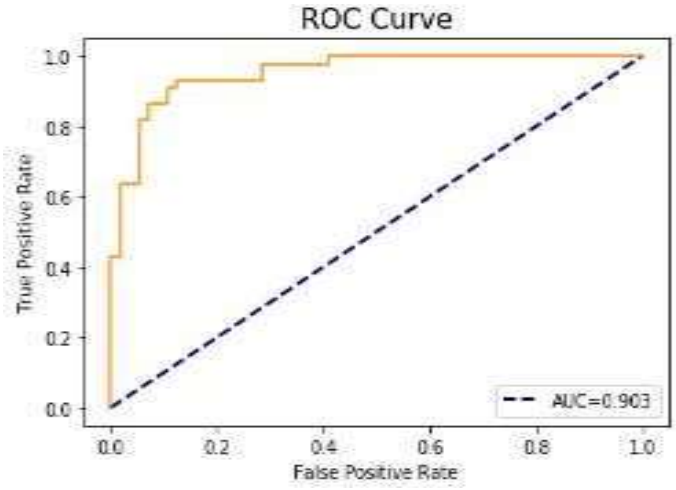


Figure 8 ROC curve of Voting

```
fpr, tpr, thresholds = metrics.roc_curve(y_test, predictions, pos_la
accuracy=accuracy_score(y_test, predictions)
precision=precision_score(y_test, predictions)
recall=recall_score(y_test, predictions)
f1 = f1_score(y_test, predictions)
auc=metrics.auc(fpr, tpr)
print("Linear SVM AUC: {:.4f}".format(auc))
print("Accuracy  : {:.4f}%".format(accuracy*100))
print("Precision : {:.4f}%".format(precision*100))
print("Recall    : {:.4f}%".format(recall*100))
print("F1-measure: {:.4f}%".format(f1*100))
```

```
Test Confusion Matrix :
[[47  4]
 [ 5 44]]
Linear SVM AUC: 0.9068
Accuracy  : 91.0000%
Precision : 91.6667%
Recall    : 89.7959%
F1-measure: 90.7216%
```

Figure 9 Result visualization of LinSVC

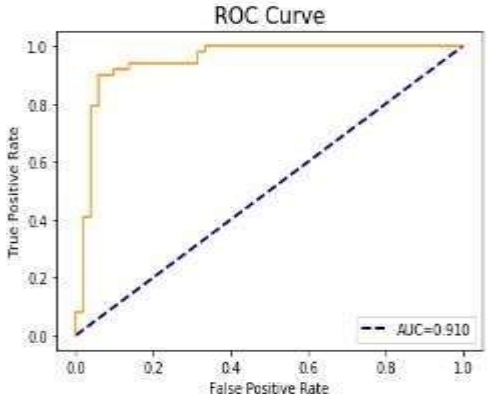


Figure 10 ROC curve of LinSVC

Table 1 showsperformance measurement of all classification algos which clearly emphasizes the performance of all algos.

Table 1. Comparison among NB, Voting, and LinSVC

Algorithm/ Performance measure	Naïve Bayes	Voting	LinearSVC
Accuracy	88.00%	90.00%	91.00%
Precision	86.79%	85.41%	91.66
Recall	90.19%	93.18%	89.79
F1 measure	88.46%	89.13%	90.72%

Figure 11 indicates the LinSVC's maximum false-positive rate in a study of three classification algos. That is a percentage between the number of adverse outcomes mistakenly defined as positive (false positive) and the total amount of true negative incidents (unless the rating is considered)

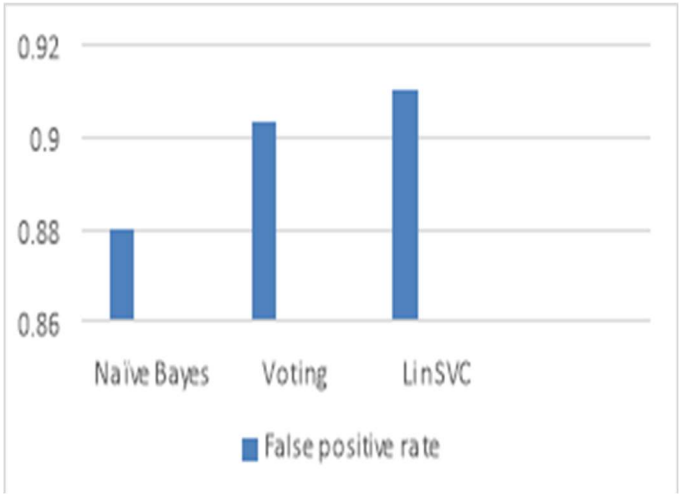


Figure 11 Comparison graph of classification algorithms forFPR

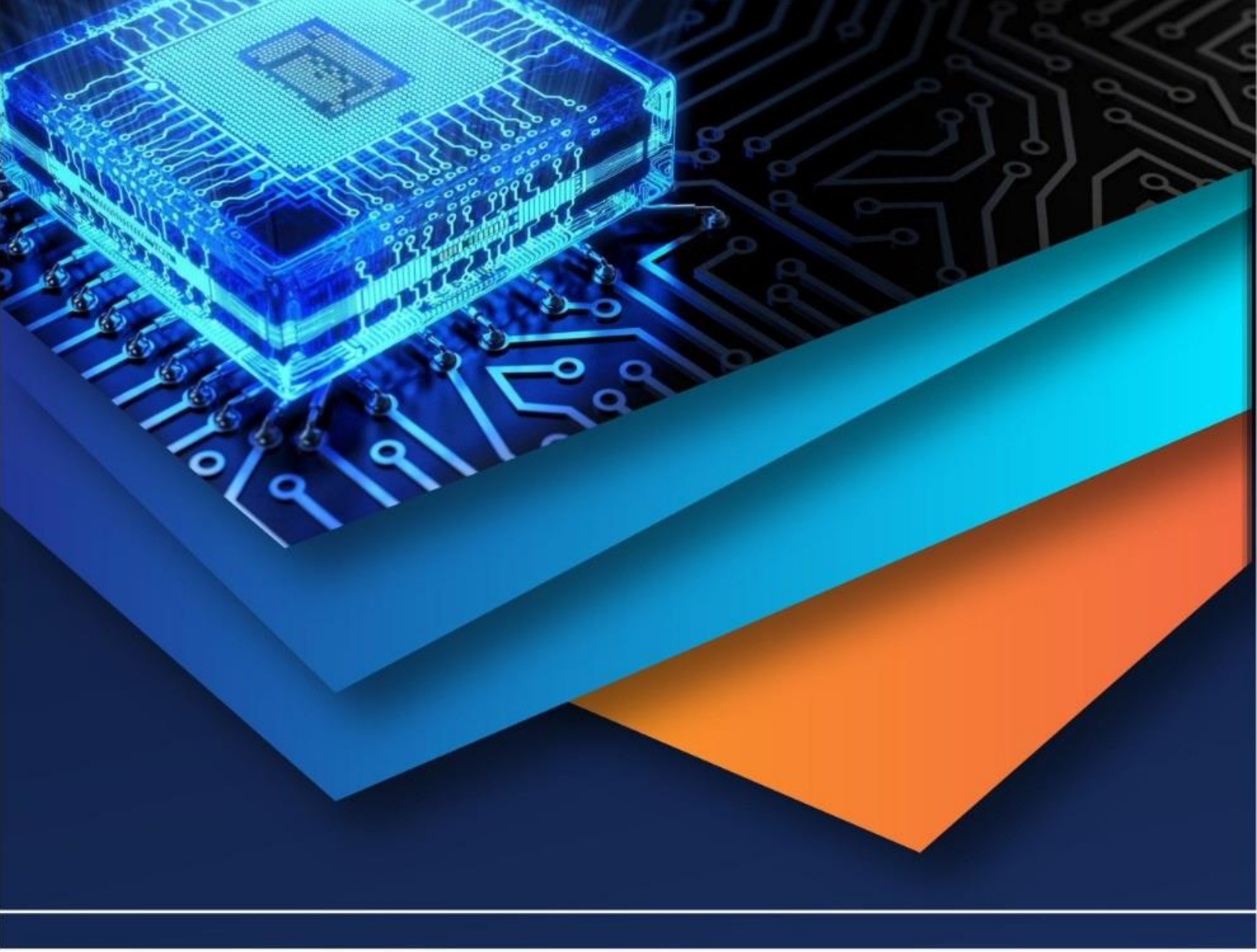
V. CONCLUSION

The primary objective of the sentimental analysis is to recognize the text according to its polarity. Opinion Mining is one of the sentimental analysis's major categories. The opinion of every consumer purchasing a product or reviewing a film relates significantly to the product or film This paper suggests an opinion analysis system for the amazon review to identify comments received from UCI Website either positivelyor negatively.The proposed approach is applied to this review dataset and obtained an accuracy of 91% by LinSVC over NBand Voting.

REFERENCES

[1] Gregory, Michelle L., et al. "User-directed sentiment analysis: Visualizing the affective content of documents." Proceedings of the Workshop on Sentiment and Subjectivity in Text. Association for Computational Linguistics, 2006.
[2] Pak, Alexander, and Patrick Paroubek. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." LREC. Vol. 10.
[3] R. Xia, C. Zong, and S. Li, "Ensemble of feature sets and classification algorithms for sentiment classification," Information Sciences, vol. 181, no. 6, pp. 1138-1152, 2011/03/15/ 2011.
[4] R. Sharma, S. Nigam, and R. Jain, "Opinion mining of movie reviews atdocument level," arXiv preprint arXiv:1408.3829, 2014.
[5] R. Sharma, S. Nigam, and R. Jain, "Polarity detection at the sentence level,"International Journal of Computer Applications, vol. 86, no. 11, 2014.
[6] Baccianella, Stefano, Andrea Esuli, andFabrizioSebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." LREC. Vol. 10. 2010.
[7] Godbole, Namrata, ManjaSrinivasaiah, and Steven Skiena. "Large-Scale Sentiment Analysis for News and Blogs." ICWSM 7 (2007): 21.
[8] Annett, Michelle, and GrzegorzKondrak. "A comparison of sentiment analysis techniques: Polarizing movie blogs." Advances in artificial intelligence. Springer Berlin Heidelberg, 2008. 25-35.
[9] Pang, Bo, Lillian Lee, and ShivakumarVaithyanathan. "Thumbs up?: sentiment classification using machine learning techniques." Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. Association for Computational Linguistics, 2002.
[10] Pang, Bo, and Lillian Lee. "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts." Proceedings of the 42nd annual meeting on Association for Computational Linguistics. Association for Computational Linguistics, 2004.
[11] C. Wu, L. Shen, and X. Wang, "A new method of using contextual information to infer the semantic orientations of context-dependent opinions," in Artificial Intelligence and ComputationalIntelligence2009. AICI'09. International Conference on, 2009, vol. 4:IEEE,pp.274-278.
[12] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, and M. Stede, "Lexiconbased methods for sentiment analysis," Computational linguistics, vol. 37, no. 2, pp. 267-307, 2011.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)