International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhancement of the Lexical Approach by N-Grams Technique via Improving Negation-Based Traditional Sentiment Analysis

Harish Dutt Sharma[1], Sanjay Sharma[2]*

Submitted: 11/12/2023 **Revised:** 23/01/2024 **Accepted:** 02/02/2024

Abstract: Sentiment analysis, often known as opinion mining, is a significant area in artificial intelligence today. Sentiment analysis was widely observed in this field. Currently, a lot of data is constantly being exchanged as text on social networking and e-commerce platforms like Facebook, Twitter, Amazon, etc. Therefore, sentiment analysis is the best technique for businesses to comprehend what their customers want from them so that they may adapt their plans in response to client feedback and expand their customer base. To extract the exact meaning from the text is a tough task. So here, our effort is to get the positive and negative sentiment of reviews from the dataset and enhance the performance of sentiment through Natural language processing (NLP) over pre-existing pre-processing technique and machine learning algorithms. So for this purpose, we have an Amazon product review dataset. Which is an extract from the Kaggle website. In this study, we aim to remove noise from the dataset and improve the traditional NLP preprocessing technique after that, we will use Term Frequency-Inverse Document Frequency (TF-IDF) method for feature selection and then classify the result through the classification algorithm such as Artificial Neural Network (ANN), Naïve Bayes (NB), and Support Vector Machine (SVM).

Keywords: *Negation handling, N-Grams techniques, Sentiment analysis, Pre-processing technique, Machine learning*

1. Introduction

Sentiment analysis (SA) fulfills the challenges and its contribution to the research area. SA deal with text, which is given in social media, where the grammar rules and spelling ignores by a person. So there we have to require a technique that pre-process the data and clean the dataset before to SA because social media has a lot of text in the form of information like fragmented sentence, emoticons, abbreviation, slang, and emojis [1]. SA is a study of customer opinion, where the unstructured data are in the form of text and unstructured data is not easy to understand because the text data have more ambiguity and irregularities as compared to formal databases. A huge amount of reviews in social media is an opportunity to get SA from the text and extract the information, which is hidden within the text that can be analyzed through the appropriate tools. So we have required such kind of tools, which analyze or visualize the document and generate the model. In such types of studies, we use several computational algorithms for visual analytics and methods that build a specific model to preprocess the reviews and classify or visualize the information as a result in sentiment analysis and extract the sentiment from customer opinion and for this entire framework, we use the NPL method for data preprocessing [2]. SA is a technique that identifies ambiguity in opinions, language etc. It is therefore also called "sentiment analysis". SA provides information about how a speaker or user thinks about a particular topic [3]. In the context of mood analysis, two concepts can be explored: 'polarity' and 'subjectivity' Subjectivity deals with a person's beliefs, opinions, or personal feelings, while polarity essentially refers to feelings expressed as negative, positive, or neutral. Sentiment analysis works at the different levels of document, sentence and sub-sentence. Various forms of sentiment analysis can be done in various areas, such as fine-grained sentiment analysis, which works with polarities ranging from very negative to very positive, or the detection of intentions or feelings. For sentence analysis, there is both a traditional lexicon-based method and a machine learning-based technique. Both methods have some pros and cons. The objective of this learning process is to identify and rank negative and positive customer feedback on different products and to use an Machine Learning (ML) model to rank them [4]. According to a study conducted by Amazon last year, more than 80% of online buyers value ratings more than explicit suggestions. Any online product with a large number of positive reviews speaks to the credibility of the item, while the lack of reviews raises doubts among potential customers. Simply put, the more reviews, the more credible they appear [5]. There are several areas where feelings analysis is used. The authors assert that sentiment analysis helps the government to identify its strengths and weaknesses by reviewing public opinion on social media [6]. In the same way, sentiment analysis is used in online commerce to identify dissatisfied customers and turn them into promoters. By analyzing their shopping experience and opinions on product quality, businesses can use sentiment analysis to make sure that customers are satisfied with their purchases. This helps businesses to build long-lasting relationships with their customers, as well as improve customer loyalty and create better customer experiences. Additionally, sentiment analysis helps to improve customer service and generate more sales and revenue. Finally, sentiment analysis also helps businesses identify trends and potential issues in their products and services, enabling them to quickly make the necessary changes to ensure customer

[1]*Department of Computer Application,*
School of CA&IT, Shri Guru Ram Rai University,
Dehradun 248001, Uttarakhand, India
ORCID: 0000-0002-6325-242X
Email: harish.research106@gmail.com
[2]*Department of Computer Application,*
School of CA&IT, Shri Guru Ram Rai University,
Dehradun 248001, Uttarakhand, India
ORCID: 0000-0001-7625-5091
Email: sanjaypokhriyal@yahoo.com
**(Corresponding Author)*

satisfaction [7]. The Author affirmed that sentiment analysis is used for assessing customer reviews and opinions about products and services in order to gain insight into customer attitudes, opinions, and emotions. The technique can also help businesses make decisions on how to improve their offerings [8]. However, it is difficult to identify, filter and monitor the feelings analysis information available on social media applications on an ongoing basis. Among these factors are the availability of unstructured data, the variety of languages, a variety of websites and social media platforms, and disparate information about people's views. As a result, suitable tools and algorithms are needed to analyze feelings from data gathered in Big Data, Blockchain, Fog Computing, and IoT-base [9]. Accurate interpretation of speech is currently almost impossible in computer programs. There are many difficulties and great challenges in the analysis of negation [10]-[12] since its contextual-dependent nature complicates bag-of-words (BoW) approaches to natural language. The latter method of document analysis considers only the frequency of words in the document, without taking into account the order of words in a document from beginning to end. However, such careful consideration is necessary, since negations occur in a variety of different forms. They reverse the meaning of single words, but also entire sentences or even phrases [13].

In this paper, we use a dataset of Amazon product reviews to obtain positive and negative sentiments. The TF-IDF technique is used to extract features, and the output is obtained using a machine learning classifier. The remainder of the work is organised as follows: in section 2, we examine the literature on opinion analysis briefly. The methods utilised are described in Section 3. Section 4 offers a thorough experimental examination of the opinions. Section 5 concludes with a conclusion and discussion of future work.

2. Literature Review

In the field of text classification and opinion analysis, much research has been done to determine the sentiment values of a text. One of the main challenges is to categorize the sentiment values of a text as either negative or positive. To address this challenge, three different forms of classification have been proposed. At the aspect level, sentiment values are classified based on individual aspects of the text. Sentence-level classification focuses on the sentiment values of an individual sentence in a text. Document-level classification takes into account the sentiment values of the entire document. These three different forms of classifying sentiment values help to accurately categorize the sentiment values of a text as either negative or positive. Worldwide, researchers have studied supervised, unsupervised, and semi-supervised machine learning techniques. The author [14] used sentiment analysis to rate reviews of the iPhone 5 on Amazon as an example. The study of reviews provided useful insights into customer experience and review sentiment. This study also provided information for future product design decisions, illustrating the importance of machine learning in today's world. This technique is a combination of various preprocessing techniques aimed at reducing noise in the data. This includes removing punctuation, HTML tags, and numbers. The part-of-speech tagger (POS) identifies the types of words in a text by identifying the individual words. After identifying the features, a rule-based procedure is used to categorize the reviews. In this procedure, a set of rules is used to assign each review to a specific category or sentiment. The author [15] conducted a study to analyze the impact of text preprocessing on online movie reviews. To reduce noise in the data, various preprocessing processes such as stemming, HTML tag removal, and data cleaning were conducted. To reduce unnecessary features, the chi-square feature selection technique was applied. After preprocessing, a support vector machine (SVM) was applied to the scores to categorize them into either negative or positive sentiments. The results of the study showed that the preprocessing steps were beneficial for classifying the ratings into the respective mood categories. The authors concluded that text preprocessing is a crucial step in the sentiment analysis of online movie reviews. The author [16] started removing noise from Amazon book reviews by applying various preprocessing techniques, such as removing URLs and HTML tags, spaces, stemming, punctuation, and special characters. This ensured that the data was ready for the next step of feature selection, which was performed using TF-IDF. TF-IDF was used to display the preprocessed data by highlighting the most important words in the reviews. This feature selection method allows the most relevant words to be identified and used in the analysis. After feature selection was completed using TF-IDF, the authors compared the accuracy of different classifiers such as K-Nearest Neighbor, Support Vector Machine, Decision Tree, Naive Bayes, and Random Forest. They also evaluated the time required by each classifier as well as the sentiment scores of different books to assess the effectiveness of each model. Unsupervised learning is a type of machine learning algorithm that works without requiring input data to be labeled or categorized. It instead relies on the data itself to identify patterns and make predictions. It is used to identify patterns and trends in data which can be used to generate insights and make predictions. Unsupervised learning can be used to estimate models without annotations, but the results are often less accurate overall [17]. Compared to unsupervised learning, supervised learning is known to perform better. For this approach to work, training sets must be created with manual labels for each word. This labeling process is very labor intensive, often requires extensive manual work, and results in only mediocre performance. In addition, this process is highly subjective, which can lead to further inaccuracies. Therefore, developing an automated method for labeling latent negation regions is critical to improving the accuracy of supervised learning [17] [18]. In the opinion analysis, the recognition of negation domains in the relevant research is mainly based on rule-based algorithms. Rule-based approaches have several disadvantages since the list of negations must be determined in advance and the selection criterion for choosing a rule is usually random or determined by cross-validation. Rules that are created with the intention of reflecting "ground truth" are inherently limited in their ability to be learned. This is because the rules are predefined based on certain assumptions and cannot be adapted to new information or experience. This means that the rules cannot evolve or grow in their understanding of the ground truth, but remain static. As a result, rules can never truly reflect "ground truth" because reality is constantly changing and evolving. Therefore, rules are not able to truly learn. Those who want to use a learning strategy can alternatively turn to generative probabilistic models [19].

In this research, we aim to address the issue of negations in reviews and feature extraction from Amazon datasets using TF-IDF. Then, to compare and assess the outcomes, we will analyse the accuracy of different classifiers as well as their sentiment scores. We will use the TF-IDF scores to study the relationships between features and ratings. We will also use the classifiers to identify the sentiment of each review, taking into account the presence of negations. Finally, we will measure the accuracy of the different classifiers and sentiment scores to determine which ones are most effective in identifying the sentiment of each review. In this way, we can better understand the impact of negations on sentiment analysis and feature extraction. The classifiers and methods used in the experiment are discussed in more detail in the following section.

3. Methodology

Most researchers have traditionally used preprocessing techniques to remove noise from reviews, as depicted in Figure 1. These techniques involve removing negative words such as "not", "no", "wouldn't", "didn't", etc. from reviews, although such negative terms can have a greater impact on sentiment, as they can change the entire meaning of a sentence. As such, it is important to consider the impact of such negative words when processing reviews. Negative terms can have a significant impact on the mood of a sentence. When we remove negative terms from a sentence using preprocessing techniques, the overall meaning and mood of the opinion can change drastically. For instance, consider the following two sentences, "He is as brave as a lion" and "A lion is not braver than him." The first sentence is a positive statement, implying that the person being referred to is very brave, just like a lion. However, the second sentence is negative because it uses the word "not" to negate the comparison between the lion's bravery and the person's bravery. The sentence implies that the lion is not as brave as the person. When we remove the negative term "not" from the second sentence, the meaning and mood change. The sentence becomes "A lion is braver than him," which is a positive statement implying that the person being referred to is less brave than a lion. Therefore, it is important to carefully consider the use of negative terms in a sentence. Negative terms can completely change the sentiment of a sentence and affect the overall meaning of a statement. Removing negative terms from documents is a common preprocessing technique that uses stopwords. While this can be useful for reducing noise in a document but it is also important to remember that it is not good practice to eliminate negative words. This can result in an incomplete or inaccurate representation of the original document, which can have a negative impact on overall accuracy. Negative terms, then, should never be neglected in a sentence. They go a long way toward ensuring that the message of the sentence is conveyed clearly and correctly.

In Figure 2, we attempt to remove negative terms from a list of stopwords. The remaining stopwords are then manually removed from the documents to ensure that no negative terms are removed from the reviews. In this way, we can deal with both negative and positive phrases in a document.

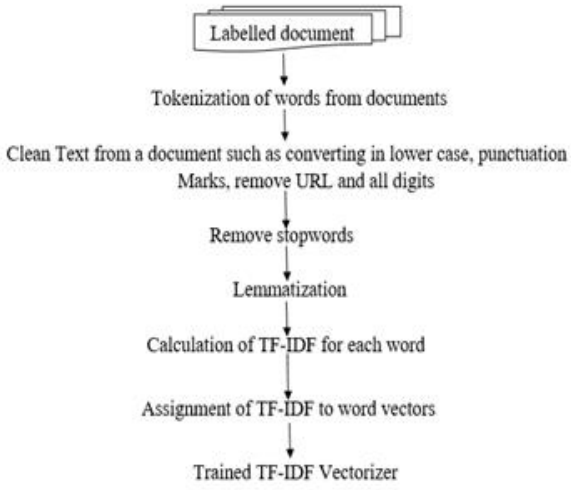

Fig. 1. The Traditional approach

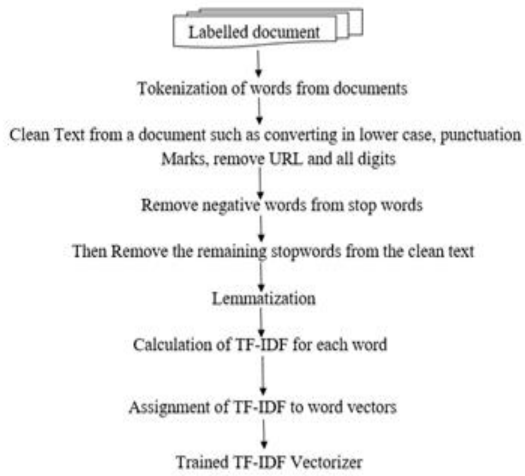

Fig. 2. Modification in Traditional approach

3.1 Data Acquisition and Pre-processing

To validate our approach, we selected the Amazon product review dataset. This dataset comprises 21,000 reviews of 30 different items, each having its own DOC_ID, LABEL, RATING, VERIFIED_PURCHASE, PRODUCT_CATEGORY, PRODUCT_ID, PRODUCT_TITLE, REVIEW_TITLE, and REVIEW_TEXT. By analyzing this dataset, we can better understand customers' experiences with different products and brands in the market. We believe this dataset provides us with enough data to accurately assess the validity of our approach. To ensure the accuracy of our dataset, we have decided to remove reviews with a 3-star rating. These reviews are considered neutral, meaning they are neither negative nor positive, and therefore need not be included. Reviews with more than three stars are considered favourable, while reviews with fewer than three stars are considered bad. We move on to the following phase with the remaining reviews, where only reviews with more than three stars are assigned a 1, while reviews with fewer than three stars are assigned a 0. After that, we clean the column "REVIEW _ TEXT" from interfering factors like URL links, HTML tags, numbers, special characters, etc. (see Figure3), and convert it into a column "Clean_Text".

	REVIEW_TEXT	Clean_Text
0	When least you think so, this product will sav...	when least you think so this product will save...
1	Lithium batteries are something new introduced...	lithium batteries are something new introduced...
2	I purchased this swing for my baby. She is 6 m...	i purchased this swing for my baby she is mon...
3	I was looking for an inexpensive desk calcolat...	i was looking for an inexpensive desk calcolat...
4	I only use it twice a week and the results are...	i only use it twice a week and the results are...
5	I'm not sure what this is supposed to be but I...	i'm not sure what this is supposed to be but i...
6	Pleased with ping pong table. 11 year old and ...	pleased with ping pong table year old and ye...

Fig. 3. Screenshot of Text cleaning

In the first step, without editing the stopwords, we remove them from the "Clean_Text" column and the text has been lemmatized (see Figure 1). The result is stored in a new column called "Before_edit_remove_sw" (see Figure 4). This process is useful for streamlining and normalizing the text in our analysis. After removing the stopwords, the text is more concise and easier to understand. Lemmatizing the text allows us to group words with the same root, which makes the analysis and identification of important terms in the data more efficient.

In the next step, stopwords are again removed from the "Clean_Text" column after negative words have been removed from the stopwords list and the text has been lemmatized (see Figure 2). The result is stored in a new column called "After_edit_remove_sw". The result of this work can be seen in Figure 4. In this step, we chose not to remove any negative words that the client may have expressed. We chose to keep the client's opinion without any influence or manipulation. This decision was made to ensure that the customer's opinion was accurately reflected in the document.

Before_edit_remove_SW	After_edit_remove_SW
least think product save day keep around case ...	least think product save day keep around case ...
lithium batteries something new introduced mar...	lithium batteries something new introduced mar...
purchased swing baby months pretty much grown ...	purchased swing baby months pretty much grown ...
looking inexpensive desk calcolatur works ever...	looking inexpensive desk calcolatur works ever...
use twice week results great used teeth whiten...	use twice week results great used teeth whiten...
'm sure supposed would recommend little resear...	not sure supposed would recommend little resea...
pleased ping pong table year old year old blas...	pleased ping pong table year old year old blas...

Fig. 4. Screenshot of dataset before edit stopwords and after edit remove stopwords

3.2 Feature Extraction

Representation is a key stage in sentiment classification. Often, noise in raw data must be filtered out using various pre-processing procedures. The preprocessed data is used to generate a term document matrix, or TDM, which specifies the frequency of each specific word. The TDM supports two feature extraction methods: the TF-IDF and the bag of words. A word's TF-IDF score is derived by multiplying the TF and IDF together; this score can also be determined by the product of these two factors. In the TF score, the terms with the highest frequency in the review dataset are weighted more heavily than other terms. The IDF scaling factor increases the relevance of the dataset's least common terms. When compared to other word categories, this score is lower for rare and common words. We are likely to eliminate terms with low TF-IDF scores if we ignore them [20].

Information can be recovered from documents using the TF-IDF technique, which takes into account both the inverse document frequency (IDF) and the term frequency (TF). Each phrase and word is assigned a distinct TF and IDF mark. The aggregate TF and IDF product scores of a phrase are added together to calculate its TF*IDF weight. To put it another way, a phrase is regarded to be rarer if it's TF*IDF score (weight) is higher, and vice versa. The TF of a term indicates how frequently it is used. The IDF of a word determines how important it is throughout the narrative. [21].

3.3 Data Representation

In this section, we present a word cloud of reviews for our Amazon product dataset (see Figure 5). A word cloud is a graphical representation of the frequency of words used in a text. This visual representation allows us to get a quick overview of the most frequently used words in the review data. By analyzing this word cloud, we can gain valuable insight into the opinions of customers who have submitted reviews. This visual representation allows us to better understand the sentiment of the reviews as well as the characteristics that customers are most interested in. This word cloud helps us better understand the preferences of the customer, which may be utilized to guide future marketing decisions.

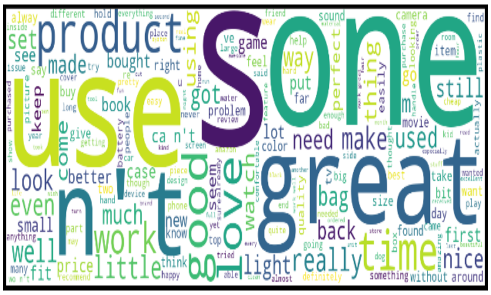

Fig. 5. Word cloud of the dataset

After conducting TF-IDF vectorization on the text in our dataset, we utilize scatter plots to display the data as either negative or positive. In order to visualize correlations between two variables, scatter plots are a valuable tool. The TF-IDF scores for each word in the text and whether or not the emotion is good or negative are the two variables in this situation. When representing the importance of a word in a text corpus for information retrieval and natural language processing, TF-IDF is frequently utilized.

We can plot the results of TF-IDF vectorization on a scatter plot once the text has been processed. The TF-IDF scores for each word in the text are represented by the x0, and the sentiment (positive or negative) is indicated by the x1. Figure 6 shows the dataset's outcome following the computation of the TF-IDF. The results of the TF-IDF method used on the dataset are shown in this figure.

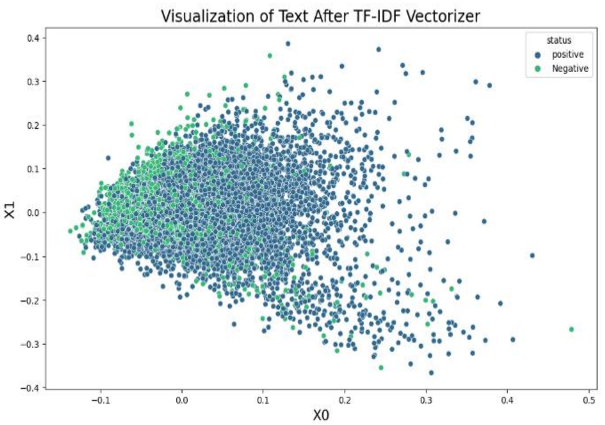

Fig. 6. Visualization of text

3.4 Experiment

We next calculated the sentiment analysis accuracy after pre-processing our dataset with two distinct techniques (as depicted in Figures 1 and 2). We achieved this by integrating the TF-IDF characteristic with unigram, bigram, and trigram. We divided the TF-IDF vectorizer into training and testing sets once we acquired it. Then, classifiers like SVM, ANN, and NB were fed these sets. When we compared the outcomes of various classifiers, we discovered that the accuracy had increased. The use of the TF-IDF feature, which allowed for a more accurate analysis of the sentiment within the dataset, is responsible for the improvement in accuracy [22]. For the suggested work, the following algorithm was used:

Proposed Algorithm

Input : labeled information
Output : Classifiers' accuracy

1. Load the dataset and preprocess it using techniques shown in Figure 1 and Figure 2.
2. Create a TF-IDF vector for each document in the dataset along with unigram, bigram, and trigram.
3. Split the dataset into training and testing sets.
4. Train the classifiers (SVM, ANN, and Naive Bayes) on the training set.
5. Test the classifiers on the testing set and calculate the accuracy of sentiment analysis.
6. Compare the accuracy of each classifier based on unigram, bigram, and trigram.
7. End

4. Results and Discussion

In this study, we present the technique and algorithm that have been recommended for evaluation. The effectiveness of the suggested technique and algorithm in handling the current issue was evaluated through a series of trials. A dataset of Amazon customer reviews that had been randomly divided into training and testing groups was used for the tests. Developing a classification model is a critical task that requires careful consideration of the features present in the dataset. As noted by author [23], the first step in developing such a model is to identify the relevant features. During model training, the words from a review can be decoded and added to the feature vector. This method is known as a "Uni-gram" if only one word is considered, and a "Bi-gram" if only two words are considered, and a "Tri-gram" if only three words are taken into account. Accurate analysis of substantial amounts of text data is made possible by combining Uni-gram and Bi-gram approaches. The accuracy of sentiment analysis can actually be greatly increased by combining these strategies, according to studies [24], [25]. Therefore, when conducting text analysis, it is strongly advised to combine Uni-gram and Bi-gram approaches.

Using a number of classifiers such as NB, SVM and ANN. we estimated the sentiment of these reviews and looked at how accurate our classifiers are for texts that can be found in the real world. Accuracy, recall, precision, and F1-score are only a few of the evaluation criteria that are used in the proposed study [26], which is described below:

4.1 Accuracy

The accuracy measure is an essential metric used in data analysis to determine how many data values have been correctly predicted. The formula for computing accuracy is as follows:

$$\text{Accuracy } = \frac{\text{Sum of Correct Predictions}}{\text{Total Predictions}} \qquad (1)$$

4.2 Precision

Among all the anticipated positive class samples, the precision measure calculates the number of samples that are actually positive as follows:

$$\text{Precision} = \frac{\text{Sumof True Positives}}{\text{Sum of True Positives + Sum of False Positives}} \qquad (2)$$

4.3 F1-score

The F1-score is a statistical measure used in machine learning to assess a classification model's accuracy. It's the harmonic mean of Precision and Sensitivity. It is also referred to as the Dice Similarity Coefficient or the Sorensen-Dice Coefficient. The perfect number is 1. F1-score computation is displayed as follows:

$$\text{F1} - \text{score} = 2 * = \frac{\text{Precision} * \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \qquad (3)$$

4.4 Recall

Recall, also known as sensitivity, is a metric used in machine learning to evaluate the performance of a classification model. It measures how many positive samples are correctly predicted by the model among all the positive samples in the test set. The formula for computing Sensitivity is as follows:

$$\text{Recall} = \frac{\text{Sum of True Positives}}{\text{Sum of true Positives + Sum of False Negatives}} \quad (4)$$

A two-phase comparative sentiment analysis is presented in this document. Figure1 illustrates the first phase, whereas Figure 2 shows the second phase. In first phase, we findings our research on the Amazon product review dataset. Table 1 displays the results of our analysis, and we did not make any modifications to the stopwords in the dataset. To preprocess the data, we followed the traditional technique shown in Figure 1. Upon examining Table 1, it is clear that ANN outperforms the other methods and provides the best results.

The results obtained from the implementation of three different machine learning algorithms, namely ANN, SVM, and NB, were evaluated based on their accuracy, precision, recall, and F1 score. The highest accuracy achieved by ANN was 90.97%, which was based on Bigram analysis. SVM, on the other hand, obtained an accuracy of 90.85% using Unigram analysis. In the case of NB, the accuracy was 84.42% using Bigram and Trigram analysis. Table 1 displays the precision, recall, and F1 score for all three algorithms.

Table 1. Classification performance prior to stopword editing

Performance measures	Unigram			Bigram			Trigram		
	NB	SVM	ANN	NB	SVM	ANN	NB	SVM	ANN
Classification accuracy	84.39	90.85	88.37	84.42	89.94	90.97	84.42	88.51	90.71
Precision	84.41	91.5	92.36	84.42	89.85	92.49	84.42	88.19	94.58
Recall	99.95	98.28	93.87	100	99.29	97.19	100	99.75	94.4
F1-score	91.53	94.77	93.11	91.55	94.33	94.78	91.55	93.61	94.49

Furthermore, Figure 7 visualizes the accuracy results of the three algorithms. It can be seen that ANN and SVM achieved higher accuracy than N.

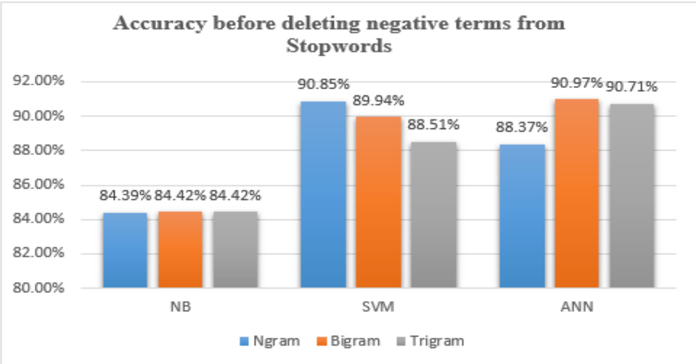

Fig. 7. Visualization of accuracy

The second phase of the procedure was preparing the data using the method depicted in Figure 2. Which involves removing negative words from the stopwords is to identify which words are considered negative. Common negative words include "no," "not," "never," "don't," and "can't". Once these words have been identified, they can be removed from the list of stopwords that are commonly excluded from text analysis because sentiment analysis is more impacted by the negative words.

The outcomes attained Based on their F1 score, recall, precision, and accuracy, the second phase of the installation of three different machine learning algorithms, namely ANN, SVM, and NB, were assessed. Based on Trigram analysis, ANN had an accuracy of 92.31%, which was its highest. On the other hand, SVM used Bigram analysis to get an accuracy of 90.88%. Using Bigram and Trigram analysis, the accuracy for NB was 84.42%, as shown in Table 2. The precision, recall, and F1 score for each of the three algorithms are shown in Table 2.

Table 2. Classification performance after stopword editing

Performance measures	**Unigram**			**Bigram**			**Trigram**		
	NB	**SVM**	**ANN**	**NB**	**SVM**	**ANN**	**NB**	**SVM**	**ANN**
Classification accuracy	84.39	90.87	84.96	84.42	90.88	91.58	84.42	89.49	92.31
Precision	84.41	91.94	90.13	84.42	90.82	92	84.42	89.25	94.41
Recall	99.95	97.75	90.33	100	99.23	98.59	100	99.54	96.61
F1-score	91.53	94.76	90.23	91.55	94.84	95.18	91.55	94.11	95.50

Figure 8 also displays the accuracy results from the three approaches. It can be seen that ANN and SVM performed more accurately than NB.

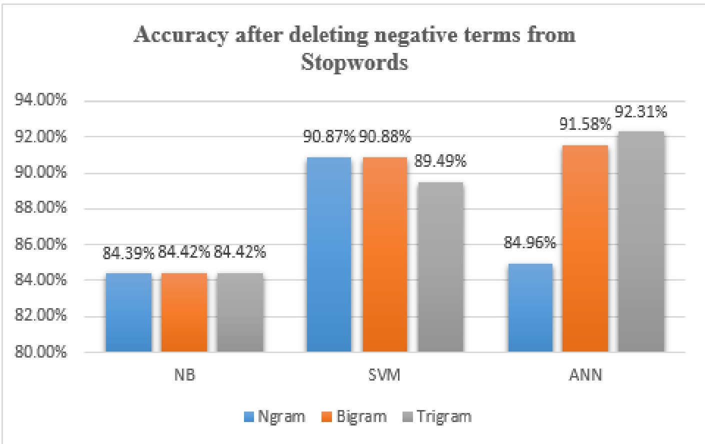

Fig. 8. Visualization of accuracy

5. Conclusion

In summary, the use of preprocessing techniques and the TF-IDF feature in combination with classifiers such as SVM, ANN, and NB proved effective in achieving higher accuracy rates. In comparison, the ANN classifier achieved the best accuracy rate of 92.31%, while the SVM classifier also performed well. However, the NB classifier did not perform well in terms of accuracy. By comparing the performance of our proposed algorithm with traditional algorithm, we showed that our algorithm was able to identify negative words more accurately. This suggests that negative words are more important than other words in the dataset. Using stopwords in NLP to eliminate unfavorable keywords from evaluations may not be the best strategy. Negative words have a big impact on sentiment analysis, thus keeping consumer feedback authentic is crucial. The proposed algorithm demonstrates that greater accuracy can be attained by leaving out negative words. In order to preserve the integrity of customer feedback, it is critical for researchers and practitioners in the field of NLP to rethink the conventional advice to delete negative terms.

Overall, future work in this sector will require a continuous development of the suggested method, which will include negative terms in our database. So we can obtain the accuracy of customer feedback without eliminating negative terms from the dataset since negative words have sentiment in sentences. We can also employ strategies such as rule-based negation handling, unsupervised procedures, and ensemble techniques to improve sentiment. These methodologies can be used to increase sentiment analysis accuracy.

Acknowledgements

Professor (Dr.) Sanjay Sharma, who served as my research adviser, has my heartfelt appreciation for his great guidance and assistance. I really thank them for everything they have done for me and the time they have spent getting to know me.

References

[1] M. A. Palomino and F. Aider, "Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis," *Appl. Sci.*, vol. 12, no. 17, Sep. 2022, doi: 10.3390/app12178765.

[2] V. Chang, L. Liu, Q. Xu, T. Li, and C. H. Hsu, "An improved model for sentiment analysis on luxury hotel review," in *Expert Systems*, John Wiley and Sons Inc, Feb. 2023. doi: 10.1111/exsy.12580.

[3] A. Iqbal, R. Amin, J. Iqbal, R. Alroobaea, A. Binmahfoudh, and M. Hussain, "Sentiment Analysis of Consumer Reviews Using Deep Learning," *Sustain.*, vol. 14, no. 17, Sep. 2022, doi: 10.3390/su141710844.

[4] K. Gupta, N. Jiwani, and P. Whig, "Effectiveness of Machine Learning in Detecting Early-Stage Leukemia," 2023, pp. 461–472. doi: 10.1007/978-981-19-2535-1_34.

[5] R. Prabowo and M. Thelwall, "Sentiment analysis: A combined approach," *J. Informetr.*, vol. 3, no. 2, pp. 143–157, Apr. 2009, doi: 10.1016/j.joi.2009.01.003.

[6] Y. Yu, W. Duan, and Q. Cao, "The impact of social and conventional media on firm equity value: A sentiment analysis approach," *Decis. Support Syst.*, vol. 55, no. 4, pp. 919–926, Nov. 2013, doi: 10.1016/j.dss.2012.12.028.

[7] R. Kumar Behera, S. Kumar Rath, S. Misra, R. Damaševičius, and R. Maskeliūnas, "Distributed Centrality Analysis of Social Network Data Using MapReduce," *Algorithms*, vol. 12, no. 8, p. 161, Aug. 2019, doi: 10.3390/a12080161.

[8] S. Vohra and J. Teraiya, "Applications and Challenges for Sentiment Analysis : A Survey," *Int. J. Eng. Res. Technol.*, vol. 2, 2013.

[9] A. Alsayat, "Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 2499–2511, Feb. 2022, doi: 10.1007/s13369-021-06227-w.

[10] N. P. Cruz, M. Taboada, and R. Mitkov, "A machine-learning approach to negation and speculation detection for sentiment analysis," *J. Assoc. Inf. Sci. Technol.*, vol. 67, no. 9, pp. 2118–2136, Sep. 2016, doi: 10.1002/asi.23533.

[11] J. Serrano-Guerrero, J. A. Olivas, F. P. Romero, and E. Herrera-Viedma, "Sentiment analysis: A review and comparative analysis of web services," *Inf. Sci. (Ny).*, vol. 311, pp. 18–38, Aug. 2015, doi: 10.1016/j.ins.2015.03.040.

[12] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," *Found. Trends® Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008, doi: 10.1561/1500000011.

[13] I. Councill, R. Mcdonald, and L. Velikovich, "What's great and what's not: Learning to classify the scope of negation for improved sentiment analysis," pp. 51–59, Jun. 2010.

[14] Aashutosh Bhatt, Ankit Patel, Harsh Chheda, and Kiran Gawande, "Amazon Review Classification and Sentiment Analysis ," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6 (6), pp. 5107–5110, 2015.

[15] E. Haddi, X. Liu, and Y. Shi, "The Role of Text Pre-processing in Sentiment Analysis," *Procedia Comput. Sci.*, vol. 17, pp. 26–32, 2013, doi: 10.1016/j.procs.2013.05.005.

[16] K. S. Srujan, S. S. Nikhil, H. Raghav Rao, K. Karthik, B. S. Harish, and H. M. Keerthi Kumar, "Classification of Amazon Book Reviews Based on Sentiment Analysis," 2018, pp. 401–411. doi: 10.1007/978-981-10-7512-4_40.

[17] N. Prollochs, S. Feuerriegel, and D. Neumann, "Enhancing Sentiment Analysis of Financial News by Detecting Negation Scopes," in *2015 48th Hawaii International Conference on System Sciences*, IEEE, Jan. 2015, pp. 959–968. doi: 10.1109/HICSS.2015.119.

[18] N. Prollochs, S. Feuerriegel, and D. Neumann, "Detecting Negation Scopes for Financial News Sentiment Using Reinforcement Learning," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, IEEE, Jan. 2016, pp. 1164–1173. doi: 10.1109/HICSS.2016.147.

[19] L. Rokach, R. Romano, and O. Maimon, "Negation recognition in medical narrative reports," *Inf. Retr. Boston.*, vol. 11, no. 6, pp. 499–538, Dec. 2008, doi: 10.1007/s10791-008-9061-0.

[20] B. Trstenjak, S. Mikac, and D. Donko, "KNN with TF-IDF based Framework for Text Categorization," *Procedia Eng.*, vol. 69, pp. 1356–1364, 2014, doi: 10.1016/j.proeng.2014.03.129.

[21] T. U. Haque, N. N. Saber, and F. M. Shah, "Sentiment analysis on large scale Amazon product reviews," in *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, IEEE, May 2018, pp. 1–6. doi: 10.1109/ICIRD.2018.8376299.

[22] B. S. Rintyarna, R. Sarno, and C. Fatichah, "Enhancing the performance of sentiment analysis task on product reviews by handling both local and global context," *Int. J. Inf. Decis. Sci.*, vol. 12, no. 1, p. 75, 2020, doi: 10.1504/IJIDS.2020.104992.

[23] A. Ritter, Mausam, O. Etzioni, and S. Clark, "Open domain event extraction from twitter," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA: ACM, Aug. 2012, pp. 1104–1112. doi: 10.1145/2339530.2339704.

[24] A. R. Razon, J. A. Barnden, and J. A. Barnden@cs, "A New Approach to Automated Text Readability Classification based on Concept Indexing with Integrated Part-of-Speech n-gram Features."

[25] M. Wankhade, A. C. S. Rao, and C. Kulkarni, “A survey on sentiment analysis methods, applications, and challenges,” *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731–5780, Oct. 2022, doi: 10.1007/s10462-022-10144-1.

[26] A. P. Rodrigues *et al.*, “Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/5211949.