Chinese Journal of Geotechnical Engineering

Volume 43 Issue 11 Year 2021

Feature Generation and Aspect Identification for Sentiment Analysis

Akshat Agarwal[*1], Ajay Singh[2], Jyoti Badola[3], Gaurav Chaturvedi[4] & Swati Dhiman[5]
[*1,2,3,4&5]Roorkee Institute of Technology, 8th K.M. Dehradun Road, Puhana, Roorkee, Haridwar (Uttarakhand) 247667

Keywords: NLP; Machine Learning; ABSA;SpaCY. Aspect Based Sentiment Analysis.

DOI:
10.11779/CJGE2021.43.3

ABSTRACT: Sentiment analysis is a technique for analyzing a piece of text to determine the sentiment contained within it.This is a powerful Artificial Intelligence system with major business ramifications. Sentiment Analysis is performedby merging natural language processing(NLP) and machine learning (ML).Using basic sentiment analysis, the software can identify whether the emotion behind a piece of text is positive, negative, or neutral.We proposed a syntax-based aspect identification algorithm to identify the sentiments of reviews. The goal of this paper is to generate a feature for aspect-based sentiment analysis using the term frequency-inverse document and bag of a word, as well as to develop a model using a statistical learning approach.The dataset includes trip advisor reviews of various hotels. There are around 20,000 reviews in this dataset. Before utilizing Bow and TF-IDF to extract features, the data was cleaned and pre-processed. Training and evaluation were performed out after the classifiers were implemented. The accuracy of a classifier is measured using evaluation metrics. Out of the four classifiers used to assess accuracy, Logistic Regression has the highest accuracy in the TF-IDF. With logistic regression,the accuracy in TF-IDF was 83percent and the classification rate in Bag of Words was 80%.

1. INTRODUCTION

The study of people's attitudes and thoughts on things like products, services, and subjects, as well as their attributes, is known as sentiment analysis.Due to its relevance to practically every imaginable issue, ranging from patron things, amenities, and medical management to social and political activities, this discipline has become an active research area since 2000[4].The most important task in aspect-based recognition, extraction of aspects in sentiment analysis is increasingly being a hot topic of research[2].ABSA is a method that considers the sentiment linked with each of the aspects is determined by the phrases associated with them.The model requires feature categories and their matching aspect words to extract sentiment for each aspect from the body of text. The most common type of opinion mining is aspect-based opinion mining, which extracts features from the text. The system is trained for higher accuracy using a supervised training method. It was formerly used to instruct the machine on review trends[3].For a specific implementation, a domain-specific model can be created; however, general language models can also be employed. For example-Take thistrip advisor hotel reviews, for example: "*The location was great but the service was poor"*Because there are multiple sentiments and topics in a single sentence in this scenario, categorizing the entire evaluation as positive or negative would be wrong. Aspect-based sentiment analysis extracts and differentiates each sentence's aspect and emotion polarity. The elements, in this case, are location and service, resulting in the sentiment attribution below:
Location was great, here Location → Positive
And the service was poor, here Service → Negative

It sorts unstructured data using natural language processing and converts it into manageable data that can be analyzed. Aspects refer to the characteristics or components of a product or service, such as the user experience of a new product, the response time for a query or complaint, or the simplicity of integrating new software. With this analysis, companies can capture the intricacies of individual attributes, components, or entities, making it easier to determine what customers like and dislike. Using our method and syntactic dependency parsing, we extract aspect as well as sentiment and calculated features, by applying a machine-learning algorithm to it. For syntactic dependency, we use *the SPACY NLP* library.

2. LITERATURE REVIEW

Analysis of sentiment is a computer process of finding and categorizing opinions expressed in a text to establish whether topic, product, or another object. In a nutshell, it's the interpretation and classification of feelings.Because it provides qualitative insights, it is particularly valuable for social media monitoring. It goes beyond the number of likes or retweets.

The methods extracted the most essential components of the evaluations and define the emotions that go with them. Inthis method, theyused language processing techniques and rules to tackle a wide range of issues in sentiment analysisandproduce a summary of finding. The aspect extraction accuracy improves dramatically when implicit aspects are taken into account, according to the ascertaining. This model outperforms support vector machine-based machine learning techniques [1]. They presented their work in this study, which contained a one-of-a-kind dataset containing domain characteristics of government smart apps as well as opinion phrases. They also described the methodology used to calculate sentiment ratings for opinion terms and built the appropriate lexicons [2]. The numerous sentiment approaches are examined in this research. Part of speech, tokenization, and lemmatization is utilized in the proposed architecture for categorizing and detecting fraudulent reviews into positive and negative categories. The suggested approach appropriately worked [3]. They used a linguistic approach to determine the opinion’sphrasedecided byprior ratings of individual words and the clause's grammatical dependency structure. A subjective lexicon is used to generate SentiWordNet. Negativity is handled with caution.Scores in the output can be utilized to find the most positive and negative clauses or sentences concerning specific movie characteristics [5]. As apart of SemEval-2014 shared task participation,they presented their work on aspect term extraction and sentiment classification in this paper. The extraction method for aspect terms is based on a supervised learning algorithm in which they used various classifiers and then combined their outputs using a majority voting mechanism. Random Forest is used to classifying sentiments [6].*ABSA* is one of three basic types of analysis, in which various features of entities are used to determine opinion orientations at the granule level.The approach of ML and DL have made significant contributions to aspect-oriented sentiment analysis surveys were presented in this work [7].They evaluatedoverallWithin the multilingual SemEval 2016 venture five datasets, they evaluated the issue and sentiment category's overall performance. They demonstrated competitive results for a variety of languages (Spanish, English, French, and Dutch) and domain likes (hotels, restaurants, electronic devices) [8].This research proposed a framework for sentiment classification based on aspects that will detect aspects fast and perform classification tasks with a high level of accuracy.The system has been tested using real-world datasets and Excel experiments, and it has been implemented as a mobile app that assists travelers in locating the best restaurant or hotel in a city [9].Using an LSTM two-layered model, this study proposed a supervised aspect-based opinion mining system. The first layer predicts the qualities described in the feedback, while the second layer determines their direction (positive, negative, or neutral) [10]. The purpose of this paper was to describe a proposed method for extracting features from opinions. When it comes to obtaining online perspectives and aspects from data, some characteristics are quite significant. The acquisition of data and comparison of alternative approaches for performing aspect-level sentiment analysis in comparison to statistical methodologies have also been examined [11]. The study's main contribution is a paradigm that assesses sentiments utilizing social network data using data mining and machine learning approaches. In comparison to the other well-known classifiers, the TF-IDF approach together with Naive Bayes fared better accuracy 81.24 percenton the social networking website Twitter [12].The normalization of the evaluation technique, which incorporates the use of shared datasets, is advocated, in order to enable the quantitative evaluation of the many proposed methodologies. One of the most potential future research directions is “connotation wealthy notion centric aspect-degree analysis, that is explored [13]. The researcher suggested a dynamic method of sentiment analysis that scans social media messages in real-time and extracts users' opinions. Thisproposedapproach entails first creating such a data-driven interactive vocabulary of word polarity. Adding new tools that allow you to fine-tune the polarity of a post employing a fixed assortment of hashtags regarding a particular topic. They classified tweets relating to the 2016 US election to validate our method.Prototype tests delineated high accuracy in sleuthing classes, furthermore as their sub-classes.[14].They devised a hierarchical neural network. Our algorithm starts by building a hierarchical LSTM model to produce sentence and document representations. Following that, user and product information is examined via attention over several semantic levels. In comparison to all state-of-the-art methodologies, the experimental findings reveal that their model made considerable and consistent improvements.[18]. The suggested method uses the Latent Dirichlet allocation (LDA) scheme to extract illustrative keywords from each subject's abstract. The k-means clustering technique is used to categorize the whole set of papers into studies papers with comparable topics[19]. This work explained the two-step method (aspect and polarity classification) for which they employed conjunction with the experimental setup just for identifying a specific piece of textual information based on the positive or negative feedback received [21].

3. METHODOLOGY

This research builds a very simple aspect-based sentiment analysis system that's able to take up generic concepts and understand the sentiments around them. For aspect extraction, we used syntax-based (also known as relation-based) approaches, which discover aspects based on the syntactical relationships they are in. Here the adjectival modifier relationship among a sentiment phrase and an aspect, as in 'nice room,' where 'nice' is adjective editing the aspect 'room,' is a reasonably easy relation. For this, we used spacy, NLP library in Python. Then calculate sentiment regarding aspect. By Using a technique of machine learning a classification framework has been constructed. The text set and its labels are part of the text classification problem; however, because text cannot be used as a model, bound numbers or vectors of numbers need to be converted into these texts.

During our study‘trip advisor hotel reviews’ dataset was tailored from *kaggle.com* was used. The data is formatted as a comma-separated value file. Out of the 20k reviews, we have taken 4000 reviews from a dataset. ‘Text’ is the first column, and ‘rating’ is the second. It has a positive ratingof (4, 5) and for negative is (1, 2), with a neutral rating of (3).We useratings and reviews for our work from this dataset.Fig.1 illustrates the proposed methodology.

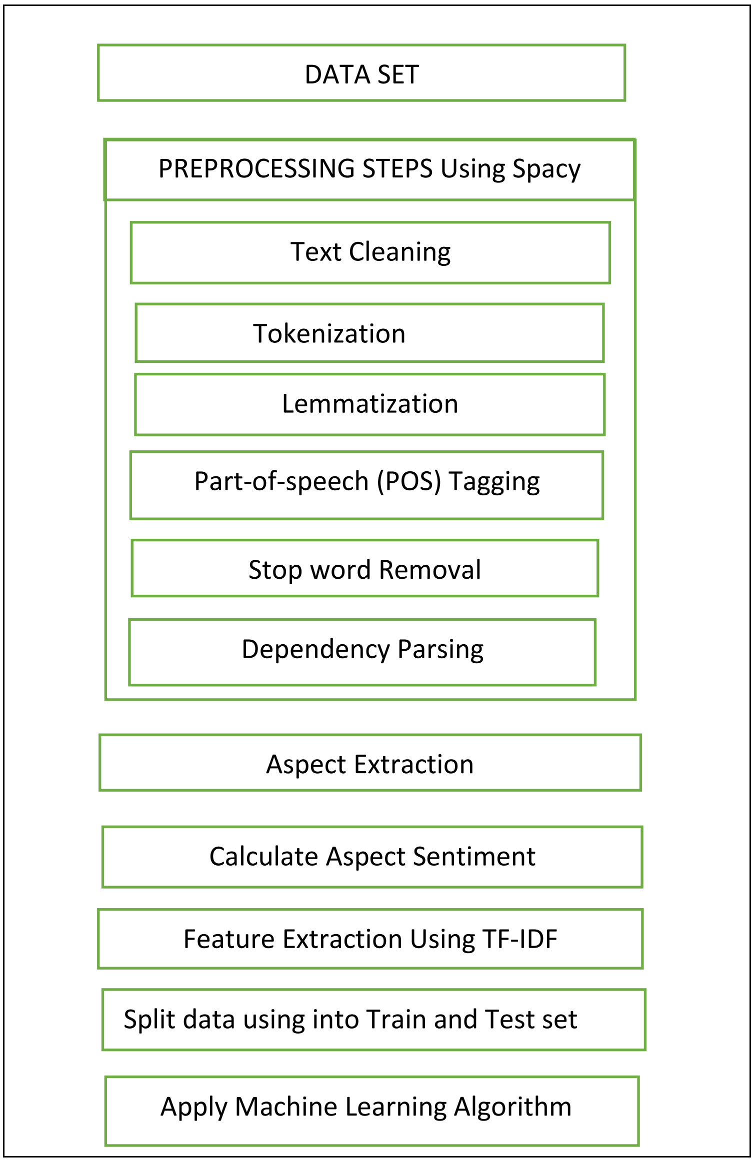

Fig1. Proposed Framework

Preprocessing Steps

Most words are uncommon, it's hard to infer raw text effectually, and it's frequent for words that appear to be wholly different to indicate virtually the same thing. When the same words are arranged in a different order, they might have completely different meanings**.** While it is possible to solve some problems using simply raw characters, it is usually preferable to complement the input with linguistic comprehension. That is precisely what spaCy is intended to accomplish. Spacy takes input in raw text and then provides the final object in the form of a doc object. We used example reviews from the dataset to demonstrate the preprocessing process. The review can be seen below-

"Nice hotel, good check fast staff friendly, breakfast cafe good......"

Our raw data was cleaned first, then tokenized, then the tokens were normalized during these operations. Tables 2 through 5 show the preprocessing results.

Text Cleaning-

At this point, any words or characters, numbers, symbols, etc. that do not contribute to the text's meaning have been removed.

a) ***Special characters removal -*** Another text preparation technique that can handle 'wonderful' and 'wonderful!' better is this one.

sub (actual pattern, replacing pattern, data)

b) ***Lowering case-*** is a basic preprocessing approach, which involves lowering the input text's unique token. To overcome the scalability problem and reduce the vocabulary size, we employed the lowercasing strategy.

c) ***Tokenization-***It is the task of breaking down a text into meaningful chunks, known as tokens. The tokenizer takes a Unicode text as input and returns a document object as output.

Table 1: Tokenization

2.2. *Review*	***2.3.*** *Tokenize word*
nice hotel good check fast staff frien dly breakfast cafe good	nice hotel good check fast staff friendly breakfast café good

Lemmatization- Words are assigned to their fundamental forms. For 'was,' the lemma is 'be,' while for 'rats,' the lemma is 'rat.

for token in doc:
print (token. Text,"--> “, token. Lemma_)

Table 2: Lemmatization

Tokens word	***After Lemma***
nice hotel good check fast staff friendly breakfast café good	nice hotel good check fast staff friendly breakfast café good

Parts-of-speech Tagging-It describes how a word works in a sentence. Nouns, pronouns, adjectives, verbs, adverbs, prepositions, conjunctions, and interjections are the eight major components of speech.

fortoken in doc:
print (token. Text, "-->", token.pos_)

Table 3: POS Tagging

Tokens word	*POS Tags*
nice	ADJ
hotel	NOUN
good	ADJ
check	NOUN
fast	VERB
staff	NOUN
friendly	ADJ
breakfast	VERB
café	NOUN
good	ADJ

Stop Word Removal-Itinclude most widely used words such as 'a', 'an', 'the', 'is', 'what' etc. When classifying or extracting information from text, these are the words that do not carry much information. As a result, we must eliminate stop words during text classification.Return 'TRUE if any stopwords are found; otherwise, return 'False'.

```
for token in doc:
print (token. Text, "-->", token.is_stop)
```

Table 4: Stop Word Removal

Tokens word	*Stopwords*
nice	False
hotel	False
good	False
check	False
fast	False
staff	False
friendly	False
breakfast	False
café	False
good	False

Dependency Parsing-*Assigning syntactic dependency labels to individual tokens, such as subject and object, to describe their relationships. Spacy includes a syntactic dependency parser that is fast and accurate. The parser also can recognize sentence boundaries and iterate over base noun phrases or chunks. Instead of tokens, chunks are used here.The several levels of the dependency parsing tree are shown in [fig.2].*

```
from spacy import display
display. render (doc, style='dep')
```

Table 5: Dependency Parsing

Token Words	*Tokened*
Nice	amod
hotel	nmod
good	amod
check	nmod
fast	amod
breakfast	compound

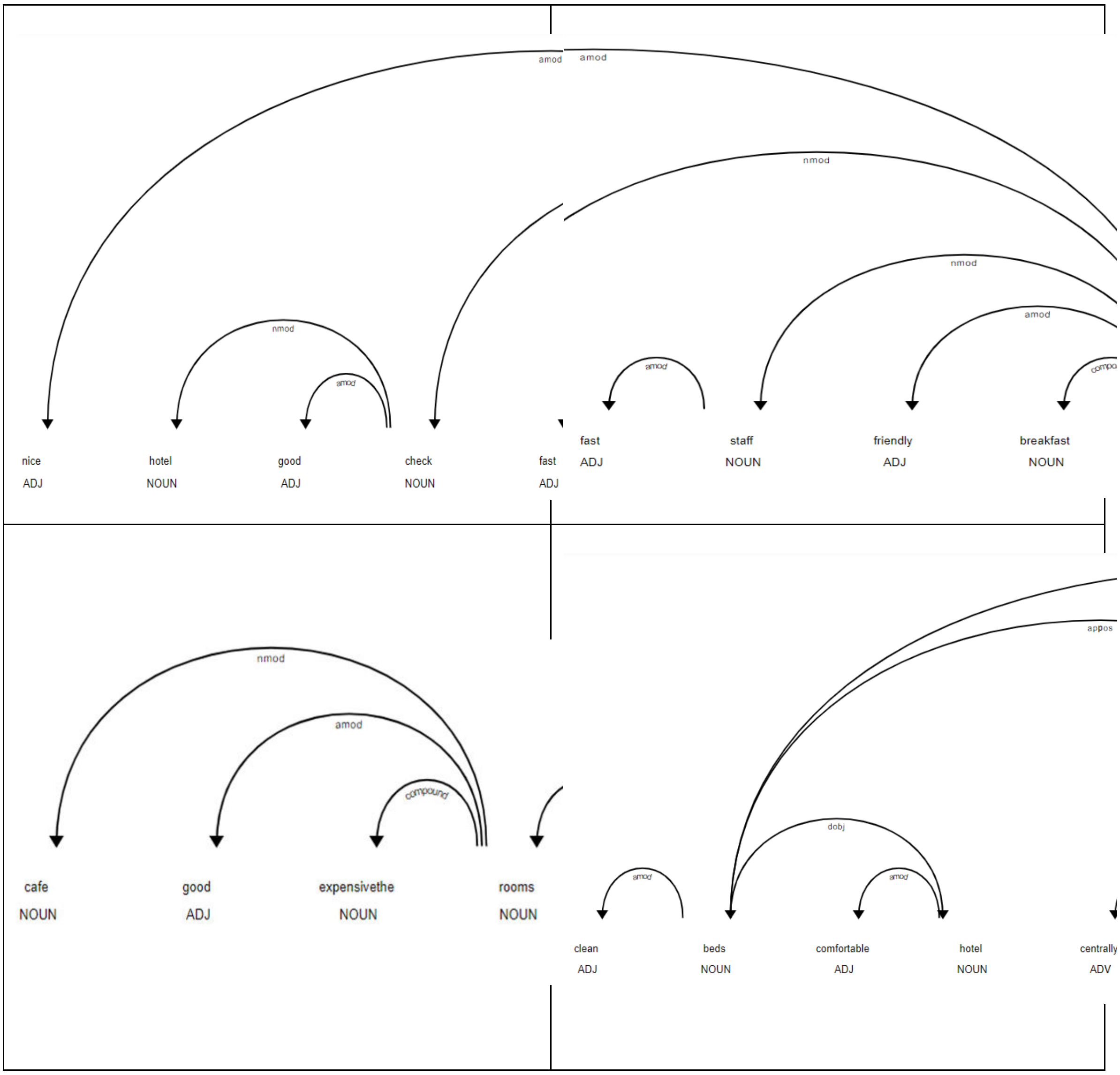

Fig.2 [Dependency Tree]

Aspect Extraction-*ABSA is a procedure that considers phrases connected with features and determines the sentiment connected with each one. For extracting the sentiment of each aspect from the concerned text, the model requires aspect categories and their related terms. If you monitor customer reviews or call transcripts, for example, you can search for aspects that have some sentiment associated with them and extract information on how to improve. We first split our sentences in such a way that the target aspects (for example, food) and their sentiment descriptions are separated (e.g., delicious). The dependency parsing and POS (Part-Of-Speech) tags may be seen for each token within our sentences. We'd also pay attention to the child tokens to see if we could locate any intensifiers. It checks for child tokens for each adjective and adverbs. We begin by extracting a sentiment description. All descriptive adjectives such as delicious, enjoyable, and tasty are picked up. It checks for child tokens for each adjective and collect adverb like "quiet," "very," and so on. We're now in a position to identify the targets mentioned.Here, we usedSyntaxbased methods. It looks for aspects based on their syntactical relationships.An adjectival modifier link between sentiment and facet is quite simple. Low-frequency aspects can be identified, which is a strong quality of these approaches. The algorithm for a related task in Fig. 3 is given below-*

Algorithm for syntax-based aspect Identification

Input: collection of sentences= {s1,s2,s3,………………,s_n}

Output: aspect assign to sentence

1. aspect[ ]
2. **for** sentence in sentences:
3. **do**
4. doc=nlp(sentence)
5. Assign space in given variables
6. descriptive_term= " "
7. target= " "
8. **for** token in doc:
9. **do**
10. **if** token.dep_=' NOUN SUB' and token.pos_= 'NOUN'
11. **then**
12. target<- token text

Fig.3

Here, we apply our approach to the dataset and show some aspect terms, descriptions,and sentiment in table 6 and 7 given below-

Table 6: Aspect and Description

Aspect	***Description***
Hotel	nice
Check	good
Staff	friendly
Café	expensive
Stay	good

Calculating aspect sentiment-It can assist us in assessing the general public's mood and emotions, as well as getting useful context information. Sentiment analysis (SA) refers to the process of analyzing data and categorizing it according to research goals. Using text blob's default sentiment analysis. It is a library that comes with built-in analysis. It uses a bag-of-words technique, which means it includes a list of words with sentiment scores linked to them, such as "good," "bad," and "great." It can also detect modifiers (like "not") and intensifiers (like "very") that have an impact on the sentiment score.

Table 7: Aspect Sentiment Description

Aspect	***Description***	***Sentiment***
Staff	friendly	Pos (0.80)
Hotel	nice	Pos (0.75)
Check	good	Pos (0.68)
Café	expensive	neg (0.40)
Stay	good	Pos (0.70)

Feature Generation using TF-IDF -Term frequency-inverse document frequency is abbreviated as tf idf. It's a method for determining the importance of words in a text by looking at how often they appear in several publications. If a term appears frequently in a text, it is significant. Assign the word a high score. A word that appears in several texts, on the other hand, isn't a unique identifier. Give the word a poor grade.

Term Frequency (tf) -It indicates the number of times each term appears in the corpus report. It's the ratio of a word's number of appearances in a report to the total number of words in that record. It raises in direct proportion to how many times that word appears in the document. There is a separate tf for each document.
Term frequency (tf) is a metric for how often text t appears in document d:

$$tf_{i,j=} \frac{n_{i,j}}{\sum_k n_{i,j}}$$

Fig 4. Formula for TF

Inverse Document Frequency (idf):*It is used to estimate the weight of unusual words across all documents. For terms that exist infrequently in the corpus, the IDF score is high. The followng isthe formula:*

$$idf(w) = \log\left(\frac{N}{df_t}\right)$$

Fig 5. The Formula For Inverse Document Frequency

Combining these two features yields the tf idf score (w) for a word in a corpus document.

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

$tf_{i,j}$ =the number of times i appears in j
N =total document number
df_i=the number of documents that containing i

Bag of Word-This model is easy to understand and apply. It's a procedure for mining text features for use in machine learning processes**.**This technique employs tokenized words to determine per token frequency for each observation.A bigram model is a method of creating a vocabulary of two-word pairs.The frequency of words is assigned within the length of the vocabulary (vocabulary refers to a collection of all the unique words).Vectorization is the process of converting text into numbers.There are two documents, for example-

Doc1: good work
Doc2: very good work

We started by defining our vocabulary, which is the collection of all words in our document set. The following are the only words discovered in the two documents:
"good, work, very"
Counting the number of occasions, a word appears in a given text, there's been created a matrix of documents and words. In Table 8, you'll see a matrix-

Table 8(Document Vector Generated)

	good	**work**	**very**	**The document's--length(in word)**
Doc1	1	1	0	2
Doc2	1	1	1	3

A bigram model, also known as a trigram model, is a combination of two or more terms, and the n-gram model is the general technique.

The general technique is referred to as the n-gram version, and a bigram or trigram model is a mixture of two or more terms. Its main advantage is its simplicity: it's cheap to compute, and often simpler is better when positioning or contextual information isn't important.

The text classifier model is built using TF-IDF and count vectorizer. Then, utilizing Logistic Regression, MultinomialNB, K-Nearest Neighbors (KNN), and Random Forest functions, import the various classifier objects. Then, using *fit ()*,We used *to predict ()* to execute prediction on test data after fitting our model to a train set.

4. RESULT AND DISCUSSION

The paper depicts how a classification model was developed utilizing a dataset with pre-defined ratings for their reviews. We created a new column named polarity from the given rating column in the dataset, with a value of zero for negative reviews, one for positive reviews, and two for neutral reviews. The data is first preprocessed during this task. To demonstrate the working of the preprocessing steps using the spacy tool, a sample review from the dataset is used. When working with spaCy, the first step is to pass a text string to an NLP object. This object is essentially a series of text pre-processing actions that the input text string must pass through. Tokenizer, tagger, parser, and many other *NLP* pipeline components are included. Cleaning and tokenization of reviews are performed we had used the parsing and tagging as well as the elimination of stopwords. After that, the data is normalized using the lemma method. It is a kind of lemmatization. Every sentence has a grammatical structure, which we extracted with the assistance of dependency parsing. It can alternatively be viewed as a directed graph, with nodes corresponding to the words in the sentence and edges between nodes corresponding to the word dependencies which is shown in fig [2] as a syntactic dependency parsing tree.Table 1-5 shows the preliminary results of sample reviews.

We created a very basic aspect-based sentiment analysis that can pick up generic concepts and analyses the sentiments surrounding them. First, we split our sentences in such a way that the target aspects (for example, location) and their sentiment descriptions are separated (e.g., good). Using spacy's dependency parsing and the POS (Part-Of-Speech) tags, we were able to see the dependency for each token within our sentences. Using child tokens, we got intensifiers like "very," "quiet," and others. Using our algorithm fig [3], we picked up all the descriptive adjectives and it checks for child tokens in each adjective and finds adverbs. Now we get aspect and their description. Using TextBlob library for sentiment calculation. In table 6-7 the intermediate outcome of aspect, their descriptive term, and sentiment are shown.

Using preprocess step join the lemma tokens and get a new review. The feature was generated using tf_idf and Bow. The training and testing sections of the dataset are separated, with test data making up 20% of the total. As a polarity, the target dataset was employed. We used numerous classifiers for the accuracy computation, including logistic regression, k-nearest neighbor (KNN), and random forest algorithms, as well as MultinomialNB for eachfeature. Objects were constructed for these functions, and then we used the fit function to outfitted the model with inside the train data and the prediction function to, are expected at the test set.

On assessment the end result said in Table 9 and 10, its miles determined that from the four classifiers used for comparing the accuracy, Logistic Regression gave the best accuracy overall performance in case TF-IDF than the BoW features. it became surpassed in the case of *TF-IDF*. The classification rate of Logistic Regression in TF-IDF was found to be**83%** whereas in the bag of words it was**80%**.

Table 9-(Tf-Idf Demonstrates Accuracy) TF-IDF

Class	***Logistic Regression***				***Random Forest***				***MultinomialNB***			
	P	***R***	***F***	***Accu.***	***P***	***R***	***F***	***Accu.***	***P***	***R***	***F***	***Accu.***
0(neg)	0.82	0.71	0.76	**0.83**	0.88	0.52	0.65	**0.79**	0.16	014	0.15	**0.58**
1(pos)	0.83	0.98	0.90		0.77	1.00	0.87		0.67	0.80	0.73	
2(neu)	0.59	0.11	0.18		**0.0**	**0.0**	**0.0**		0.24	0.04	0.07	

k-nearest neighbors (KNN)				
Class	***P***	***R***	***F***	***Accu.***
0(neg)	0.66	0.47	0.55	**0.76**
1(pos)	0.78	0.96	0.86	
2(neu)	0.38	0.03	0.06	

Table 10- (Accuracy Using Bag-Of-Words) bag-of- word

Class	**Logistic -Regression**				**Random-forest**				**MultinomialNB**			
	P	**R**	**F**	**Accu.**	**P**	**R**	**F**	**Accu.**	**P**	**R**	**F**	**Accu.**
0(neg)	0.72	0.68	0.70	**0.80**	0.96	0.17	0.29	**0.72**	0.0	0.0	0.0	**0.69**
1(pos)	0.85	0.94	0.89		0.72	1.00	0.83		0.69	1.0	0.81	
2(neu)	0.38	0.22	0.27		**0.0**	**0.0**	**0.0**		0.0	0.0	0.0	

k-nearest neighbors (KNN)				
Class	**P**	**R**	**F**	**Accu.**
0(neg)	0.86	0.13	0.23	**0.73**
1(pos)	0.74	0.98	0.84	
2(neu)	0.30	0.11	0.16	

The terminology used here is precision(p), recall(R), F-measure (F), and accuracy(Accu).

5. CONCLUSION

Sentiment analysis is one of NLP's most widely used applications. It takes advantage of the large amounts of data available on public platforms. And it gives businesses vital information that helps them improve their services and, as a result, increase consumer happiness. Sentiment analysis based on aspects is a step forward from traditional sentiment analysis. However, the simpler models can deliver acceptable results with significantly less compute and training time. We developed a classification model in this paper using a supervised technique. The text was subjected to general text analytics operations such as SPACY tokenization, normalization, stop word removal, dependency parsing, and POS tagging, etc. The syntactic parsing tree is built.In this work, the feature-generating tools TF-IDF and Bow were used. Thedatasetwas collected from *Kaggle.com*. We developed aspect and its descriptive term using our algorithm on reviews of the dataset. TextBlob was used to compute the sentiment. We need to use feature engineering to improve the accuracy so that sentiment performance can improve as well. With an accuracy of 83.0 percent, our model provides the best result. The limitations include the inability to work effectively across domains, low sentiment analysis accuracy and performance due to a lack of labeled data, and difficulty to deal with complex sentences that require more than sentiment terms and simple analysis. This may be stepped forward withinside the future through leveraging supervised statistics for a particular area and deep learning techniques for aspect-based sentiment analysis, so as to be confined to a single domain and a pre-described listing of aspect-categories.

REFERENCES

1. Alqaryouti, O., Siyam, N., Monem, A.A., Shaalan, K. (2019), “Aspect-based sentiment analysis using smart government review data” Applied Computing and Informatics, Vol. ahead-of-print No. ahead-of-print. https://10.1016/j.aci.2019.11.003. The original publication date for this paper was 23/11/2019
2. O. Alqaryouti, N. Siyam, K. Shaalan, “A Sentiment analysis lexical resource and dataset for government smart apps domain”, in A. Hassanien, M. Tolba, K. Shaalan (Eds.), Proceedings of the International Conference on Advanced Intelligent Systems and Informatics on Advanced Intelligent Systems and Informatics 2018 (AISI2018), Springer, Cham, 2019.
3. Shubham D, Mithil P, Meesala Shobharani, Sumathy S*“Aspect level sentiment analysis using machine learning”, IOP Conf. Series: Materials Science and Engineering 263 (2017) 042009 doi:10.1088/1757-899X/263/4/042009, *Email: ssumathy@vit.ac.in

4. B. Liu, Sentiment analysis and opinion mining, Synthesis Lectures Human Lang. Technol. 5 (2012) 1–167, http://dx.doi.org/10.2200/S00416ED1V01Y201204HLT016
5. Tun Thura Thet, Jin-Cheon Na and Christopher S.G. Khoo,"Aspect-based sentiment analysis of movie reviews on discussion boards", Journal of Information Science, XX (X) 2010, pp. 1–26 © CILIP, DOI: 99, May 10, 2016
6. D.K. Gupta, A. Ekbal, "Supervised MachineLearning for Aspect Based Sentiment Analysis", Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), pages 319–323, Dublin, Ireland, August 23-24, 2014.
7. Syam Mohan E, R. Sunitha," Survey on Aspect Based Sentiment Analysis Using Machine Learning Techniques", European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 07, Issue 10, 2020
8. Aitor Garc´ıa-Pablosa,∗ , Montse Cuadrosa, German Rigau," W2VLDA: Almost Unsupervised System for Aspect Based Sentiment Analysis", July 19, 2017
9. M. Afzaal, M.Usman, A.Fong, "Tourism Mobile App with Aspect-Based Sentiment ClassificationFramework for Tourist Reviews",Citation information: DOI 10.1109/TCE.2019.2908944, IEEE Transactions on Consumer Electronics
10. IRUM SINDHU, SHER MUHAMMAD DAUDPOTA, KAMAL BADAR., MAHEEN BAKHTYAR, JUNAID BABER, AND MOHAMMAD NURUNNABI, "Aspect Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation", Citation information: DOI 10.1109/ACCESS.2019.2928872, IEEE
11. Neha Nandal, Jyoti Pruthi, Amit Choudhary, "Aspect Based Sentiment Analysis Approaches with Mining of Reviews": A Comparative Study, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6, March 2019
12. S. Ahmed, S. Hina, E. Atwell, F. Ahmed, "Aspect Based Sentiment Analysis Framework using Data from Social Media Network", IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.7, July 2017
13. K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis", DOI 10.1109/TKDE.2015.2485209, IEEE Transactions on Knowledge and Data Engineering JOURNAL
14. Imane El Alaoui1,2*, Youssef Gahi3, Rochdi Messoussi1, Youness Chaabi1, Alexis Todoskof2 and Abdessamad Kobi2, "A novel adaptable approach for sentiment analysis on big social data", El Alaoui et al. J Big Data (2018) 5:12 https://doi.org/10.1186/s40537-018-0120-0
15. D.D. Dsouza, Deepika, D. P Nayak, E. J. Machado, N. D. Adesh," Sentimental Analysis of Student Feedback using Machine Learning Techniques ", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019, Retrieval Number: A11 810681S419/19©BEIESP
16. W.Zhou, W. Hanbin, S. Hongguang, and S.Tieli," A Method of Short Text Representation Based on the Feature Probability Embedded Vector",* Received: 4 July 2019; Accepted: 26 August 2019; Published: 28 August 2019
17. Anshuman, S. Rao, M. Kakkar, "A Rating Approach based on Sentiment Analysis", · January 2017 DOI: 10.1109/CONFLUENCE.2017.7943213
18. Huimin Chen, Maosong Sun1,2∗ , Cunchao Tu, Yankai Lin, Zhiyuan Liu," Neural Sentiment Classification with User and Product Attention" Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1650–1659, Austin, Texas, November 1-5, 2016. c 2016 Association for Computational Linguistics,http://dx.doi.org/10.18653/v1/D16-1171
19. Sang-Woon Kim, and Joon-Min Gil,Comput. Inf. Sci. (2019) 9:30, Research paper "classification systems based on TF-IDF and LDA schemes", https://doi.org/10.1186/s13673-019-0192-7
20. 20-Saif M. Mohammad, Challenges in Sentiment Analysis, E. Cambria et al. (eds.), "A Practical Guide to Sentiment Analysis", Socio-Affective Computing 5, DOI 10.1007/978-3-319-55394-8_4
21. Chetashri Bhadanea, Hardi Dalalb, Heenal Doshi, Sentiment analysis: Measuring opinions, Peer-review under responsibility of scientific committee of International Conference on Advanced Computing Technologies and Applications (ICACTA-2015). DOI: 10.1016/j.procs.2015.03.159.
22. M. Umair, A. Hakim, A. Hussain, and S. Naseem, "Sentiment Analysis of Students' Feedback before and after COVID-19 Pandemic", International Journal on Emerging Technologies 12(2): 177-182(2021)

23. Munir Ahmad1, Shabib Aftab2, Iftikhar Ali3, and Noureen Hameed4, Hybrid Tools and Techniques for Sentiment Analysis: A Review, INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY SCIENCES AND ENGINEERING, VOL. 8, NO. 4, JUNE 2017, [ISSN: 2045-7057]
24. R. S. Jagdale, V. S. Shirsat, S. N. Deshmukh," Sentiment Analysis on Product Reviews Using Machine Learning Techniques", Advances in Intelligent Systems and Computing 768, https://doi.org/10.1007/978-981-13-0617-4_61
25. Salud M. Jimenez-Zafra, Eugenio Mart ´ ´ınez-Camara, ´ M. Teresa Mart´ın-Valdivia, L. Alfonso Urena-L ˜ opez, SINAI: Syntactic approach for Aspect Based Sentiment Analysis, Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), pages 730–735, Denver, Colorado, June 4-5, 2015. c 2015 Association for Computational Linguistics
26. M. H. Phan & P. Ogunbona, "Modelling Context and Syntactical Features for Aspect-based Sentiment Analysis", Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 3211–3220 July 5 - 10, 2020. c 2020 Association for Computational Linguistics
27. Mr. M. Y. Babu, Dr. P. V. Pal Reddy, Dr. C. S. Bindu, COMBINED APPROACH FOR ASPECT TERM EXTRACTION IN ASPECT-BASED SENTIMENT ANALYSIS, JOURNAL OF CRITICAL REVIEWS ISSN- 2394-5125 VOL 7, ISSUE 18, 2020
28. Bhavana R. Bhamare1 and Jeyanthi Prabhu, "A supervised scheme for aspect extraction in sentiment analysis using the hybrid feature set of word dependency relations and lemmas", Bhamare and Prabhu (2021), PeerJ Comput. Sci., DOI 10.7717/peerj-cs.347
29. Kulkarni, SS & Edwards, DJ 2022, 'A bibliometric review on the implications of renewable offshore marine energy development on marine species', Aquaculture and Fisheries, vol. 7, no. 2, pp. 211-222. https://doi.org/10.1016/j.aaf.2021.10.005
30. Kulkarni, S, Wang, L & Venetsanos, D 2022, 'Managing Technology Transfer Challenges in the Renewable Energy Sector within the European Union', Wind, vol. 2, no. 1, pp. 150-174. https://doi.org/10.3390/wind2010009
31. Kulkarni, S & Chima, P 2021 'Challenges faced by UK university students due to the coronavirus crisis in the Higher Education' Preprints.org. https://doi.org/10.20944/preprints202102.0192.v1
32. Kulkarni, SS, Wang, L, Golsby, N & Lander, M 2021, 'Fluid-structure interaction based optimisation in tidal turbines: A perspective review', Journal of Ocean Engineering and Science. https://doi.org/10.1016/j.joes.2021.09.017
33. Kulkarni, S, Edwards, DJ, Chapman, C, Hosseini, MR & Owusu-Manu, D-G 2019, 'A preliminary mechanical design evaluation of the Wikispeed car: for light-weighting implications', Journal of Engineering, Design and Technology, vol. 17, no. 1, pp. 230-249. https://doi.org/10.1108/JEDT-09-2018-0154
34. Kulkarni, S, Chapman, C, Shah, H & Edwards, D 2018, 'A computational design method for bio-mimicked horizontal axis tidal turbines', International Journal of Building Pathology and Adaptation, vol. 36, no. 2, pp. 188-209. https://doi.org/10.1108/IJBPA-06-2017-0029
35. Kulkarni, S, Chapman, C, Shah, H, Parn, E & Edwards, DJ 2018, 'Designing an efficient tidal turbine blade through bio-mimicry: A systematic review', Journal of Engineering, Design and Technology, vol. 16, no. 1, pp. 101-124. https://doi.org/10.1108/JEDT-08-2017-0077
36. Kulkarni, S, Chapman, C, Shah, H, Parn, E & Edwards, DJ 2018, 'Design Study of Horizontal Axis Tidal Turbine Blade', Mindanao Journal of Science and Technology. https://www.open-access.bcu.ac.uk/10089/
37. Kulkarni, S, Edwards, DJ, Parn, E, Chapman, C, Aigbavboa, CO & Cornish, R 2018, 'Evaluation Of Vehicle Lightweighting To Reduce Greenhouse Gas Emissions With Focus On Magnesium Substitution', Journal of Engineering, Design and Technology, vol. 16, no. 6, pp. 869-888. https://doi.org/10.1108/JEDT-03-2018-0042
38. Kulkarni, S & Sodre, JR 2017, An Assessment of Lightweight Vehicles to Reduce Greenhouse Gas Emissions with Focus on Magnesium. in Thoughts and Refections on the use of Magnesium. Birmingham City University, pp. 84-91.
39. Kulkarni, S, Chapman, C & Shah, H 2016 'Computational Fluid Dynamics (CFD) Mesh Independency Study of A Straight Blade Horizontal Axis Tidal Turbine' Preprints.org. https://doi.org/10.20944/preprints201608.0008.v1.