Journal of Information Systems Engineering and Management
2025, 10(22s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Tweet Sentiment Analysis by Python and Classification by Attention Sequence Modeling

Medvedeva Marina Alexandrovna[1], Al-LAMI Mustafa Ali Mohsin[2]
1,2Ural Federal University
marmed55@yandex.ru
mustafaali1994101@gmail.com

ARTICLE INFO

Received: 20 Dec 2024

Revised: 30 Jan 2025

Accepted: 15 Feb 2025

ABSTRACT

Tweet sentiment analysis with Python and machine learning is a natural language processing task that involves determining the sentiment of tweets. In this task, the objective is to classify a given tweet as positive, negative, or neutral. Sentiment analysis can be useful in a variety of applications, such as brand reputation management, customer feedback analysis, and political analysis. To perform tweet sentiment analysis, you need to preprocess the tweet data, extract relevant features, and train a machine learning model to classify the tweets based on their sentiment scores. The most common machine learning models used for sentiment analysis include Naive Bayes, Support Vector Machines (SVM), and Recurrent Neural Networks (RNN). In this process, we collect tweet data, preprocess it, extract features and then train the model. We can use the training data to train our model and then evaluate its performance on a separate set of test data. Once the model is trained, we can use it to predict the sentiment of new, unseen tweets. Python provides several libraries such as pandas, scikit-learn, and NLTK that can be used for tweet sentiment analysis. By using these libraries, we can create a complete end-to-end pipeline for tweet sentiment analysis.

Keywords: Tweet sentiment analysis, Python, Machine Learning, Natural Language Processing.

I. INTRODUCTION

The undertaking acknowledges about the increasing use of social media and how rapidly it has been influencing the sentiment analysis of people on various matters. Though the users' opinion is different, volatile and vulnerable on each and every topic, still retrieving knowledge from the data is critical. Along with information regarding the websites they visited, understanding that the content carries important sentiments across several platforms has become essential for gauging public opinion about a certain subject. As a result, a prominent sentiment analysis technique is categorizing a text's polarity. [1].

The texts may be neutral, negative, or positive in tone, but they all generally refer to a range of emotions from happiness to melancholy. The word "sentiment" designates a topic that is both subjective and objective, as well as a topic that is both practical and imagined and blurs the line between positive or negative subjects. The strategy is grounded in text analysis since sentiment analysis depends on the propagation of myths or gossips. It involves determining the subjectivity of a belief and the outcome of a tweet; therefore, it is concerned with classifying opinion an individual into different classes according to data size and type of document. There are different schemes applied for sentiment analysis depending on a variety of methods of NLP (Natural Language Processing) and ML (Machine Learning) techniques to extract sufficient features and classify text into suitable polarity labels [2]. This chapter shows an overview of types of sentiment analysis, comparison of different techniques applied, and their importance attached with it while doing analyzing sentiments.

A. NECESSITY OF SENTIMENT ANALYSIS

Sentiment is an expression of feelings that manifest in actions, and often influence others in their behavior. Therefore, study is undertaken how to leverage these sentiments called Sentiment Analysis (SA), computationally, which determines the outcome of the result. In this study, Opinion mining is a crucial application of NLP in addressing the interplay between human and machine language [3]. Building a system to gather and organize data from various social network sites, including blogs, online discussion boards, social media, web- based surveys, is called opinion mining.

Since people take decisions based on the available resources, and to some extent social media often helps them, resulting business houses getting an opportunity to understand the buyer's attitude, appraisal, sentiment and opinion. Similarly, policymakers and politicians make adjustments while dealing with public issues. A possibility to create a new application using real-world data now exists, with a special emphasis on sorting, finding, or analyzing textual knowledge discovery techniques, specifically Twitter. Twitter is used to tweet messages via

social media and blogs sites, where new issues arise as a result of different distinctive traits contained in tweets that also influence the various the diverse areas and methodologies discussed in sentiment analysis. The key attributes of tweets are as follows, which might help you comprehend them better [4]:

- **Length of Messages**: Because messages are limited to 280 characters, reviews of products and movies, like long text classifiers, rely on particular connections to prior feelings.
- **Writing System**: In spite of unseemly spellings, cyber slangs, acronyms and unique characters the message is conveyed without changing the meaning.
- **Accessibility:** More individual tweets are publicly avail- able than on Facebook, and an API makes it easier to compile these tweets.
- **Topics:** On one side, there are a lot of existing sites focusing on a particular topic, and on the other, range of topics covered are also vast, still, Twitter is able to differentiate from its past using specific domains.
- **Real-Time**: Blogs are longer, and updated in longer intervals, which is due to paucity of time. [5]

B. *TYPES*

Analysis is a sort of research in NLP that seeks to recognize subjectivity in the text and/or extract and classify opinions and viewpoints. It investigates how people feel about various services, goods, people, organizations, subjects, events, and their characteristics. It also examines how these feelings are expressed in written and verbal form. According to the following standards, the text is categorized in sentiment analysis:

- Positivity, negativity, and neutrality of the emotions expressed;
- Result polarity, such as in medical texts when progress is contrasted with death;
- Accord or dispute on a subject, as in political conversations and debates;
- pleasant or negative news;
- In favor of or against;

C. *IMPORTANCE OF SENTIMENT ANALYSIS*

Sentiment Analysis gives incredible inputs to restructure the policy decisions or restructure organizational plans, this applies in public discourse or while being a CMD of a conglomerate, here are a few advantages. [6]:

- **Marketing Approach:** Companies use social sites as a platform to promote brands and services, and to optimize marketing strategy, continues sentiment analysis is extremely useful.
- **Enlarge Product Quality:** It is easier through marketing research to establish the quality of the product as opinion of the customer matters, and certainly it influences others demand of choices.
- **Improve Customer Service:** The customers are micro- influencers; thus, the utmost priority of a company is too pro- vide the best customer service, on-time delivery, appropriate compensation for defective products, active response on social media, etc. Under these circumstances, sentiment analysis addresses negative impact on the customers [7].
- **Crisis Management:** Since the social media influences the current issues and challenges, its constant monitoring helps mitigates unforeseen crisis, thus prevents further damages to the company or brand-name.

D. CHALLENGES OF SENTIMENT ANALYSIS

Even when NLP and Text Analysis Techniques (TAL) help in discovering and deriving the particular data, Finding the appropriate emotion polarity and substantiating the right perception of feelings are still challenges. The following are a few difficulties:

- **Sentiment of the speaker**: Although the feelings and views expressed by the users do not necessarily influence the buyers, still the customer's review of a particular product can certainly polarize the decision of the customer.
- **Multiple Thoughts**: Having multiple thoughts on services, qualities and presentation of the products pose challenging due to the absence of the same sentiments towards the same entity.
- **Fake Opinion**: It is witnessed that customers are often misguided by the fake reviews, thus the sentiment opinions are useless in various applications due to the presence of spam. [8]
- **Challenge in Application:** Though political parties, companies and film producers commonly apply sentiment analysis, still they lack competency in tracking systems which can give optimal result.
- **Language Problem**: Other than English language using Opinion Mining (OP) in other languages are difficult because of lexicons, corpora and dictionaries, therefore, a challenge for the researchers to build resources [9]
- **Sarcasm Detection**: People use positive words to convey their negative opinions in sarcastic text. This element enables sarcasm for cheating sentiment analysis models effortlessly except they are particularly devised to consider this likeli- hood [10]. Sarcasm generally occurs in contents created by users including Facebook comments, tweeter messages, etc. Also, it is not an easy task to detect Sarcasm detection in sentiment analysis without adequate comprehension regarding the circumstances, the particular topic, and the situation.
- **Detecting negativity:** In semantics, negation is a technique to flip the polarity of words, phrases, and sentences. Researchers utilize a variety of semantic rules to determine whether negation is present. On the other hand, it is crucial to establish the scope of the words that negation words influence [11].
- **Word Ambiguity**: Uncertainty in words is another problem with sentiment classification. The problem of word ambiguity means that it is impossible to define polarity beforehand because it varies so strongly on the context of the phrase for particular words.
- **Multi-polarity**: Sometimes a sentence, document, or other text unit used for analysis displays multi-polarity. In certain situations, relying solely on the analysis's overall findings could be deceptive, much like how an average occasionally hides crucial details about all the inserted statistics.

E. CHALLENGES OF SENTIMENT ANALYSIS

Following are the various essential features and characteristics of the sentiment analysis: -

- **Contributor**: These are the individuals or groups who express their thoughts in written or textual form. A contributor may also contact the source or the opinion bearer.
- **Object**: A commodity, service, people, incident, organization, or subject can all be considered objects [12]. It might be connected to a group of elements and qualities that are referred to as aspects in sentiment classification studies.
- **Review**: A review is a text created by a contributor that expresses the contributor's ideas regarding a certain element of the object. Another name for a review is an opinionated paper.
- **Overall rating**: For instance, on a Likert scale from 1 to 5, this is a user-reported overall satisfaction with the product. An opinion on a subject is a contributor's favorable, neutral, or negative attitude, feeling, or assessment of that subject.
- **Aspect**: According to the contributor's comments in their review, an element is a crucial component of the product in terms of total service quality [13].

F. *SENTIMENT ANALYSIS LEVELS*

Sentiment analysis can take place at several levels, including the document, sentence, and aspect levels. Below is a quick summary of each of these levels:

1) *Document Level:* Once sentiment is retrieved from the full review, a whole opinion is categorized using document level analysis based on the reviewer's overall sentiment. This level's main goal is to categorize documents that communicate opinions, whether they are good or negative. For various publications, the approaches utilized for this approach are 70 to 80% accurate [14]. Reviews of products and films are its principal uses. Sentiment analysis at the document level only considers one entity. This technique is not appropriate for document evaluation or comparison of many items.

2) *Sentence Level:* Determine if a statement is subjective or objective as the first step in this level. The former refers to the individual's interpretation, while the latter denotes that the reader is viewing the sentence from a distance or from the perspective of another person [15]. The major objective of the second task is to determine if the subjective sentence is positive, negative, or neutral. This technique mostly consists of two steps:

- The subjective classification of a sentence can be broken down into two categories: objective and subjective;
- Positive and negative sentiments are classified as subjective sentences.

A subjective statement expresses unique feelings, views, sentiments, or values, as opposed to an objective sentence, which typically conveys accurate information. The Naive Bayesian Classifier approach is used to distinguish between the two. However, it is not enough to know if the sentence is expressing a positive or negative attitude. [16]. This phase in the middle is beneficial to filter out the sentences that don't express any viewpoints, yet a subjective statement can contain a variety of opinions, along with subjective and accurate components.

3) *Aspect Level:* When performing fine-grained analysis, the aspect level categorization, also known as entity, feature, or phase level, emphasizes the viewpoints rather than focusing on the linguistic construct (i.e., document, paragraphs, sentence, clauses & phrases). By considering the positive and negative emotions of a review, it enables the discovery of object sentiments and its features. Next, it detects and extracts the features of an object expressed by the reviewer. After the synonym grouping is completed, the feature-based feedback summary is generated. Aspect extraction and aspect emotion classification are its two main functions [17].

- The first objective is to identify an object's characteristics, which is frequently referred to as an information retrieval function.
- The second assignment establishes the viewpoints' degree of positivity, negativity, or neutrality.

G. *USE OF SENTIMENT ANALYSIS*

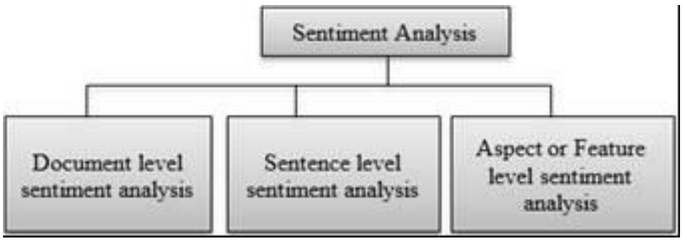

Figure 1. Sentiment Analysis Levels

It is experienced that when buying a product, people often seek opinion of others. It is to establish the credibility of the specific features of a product. And, depending upon the intelligence and wisdom of the person, one would certainly not depend on a single global rating. Therefore, the first application of sentiment analysis gives an indication based on choice [18]. The second one is, that the companies seek input of the customers, and, the last one is through the application of technology which enables text analysis automatically. There are some other applications listed below:

- **E-Commerce**: Almost all e-commerce activities make use of sentiment analysis. Every electronic commerce site permits clients to give their reviews regarding shopping and product quality. The clients give scores and rating to goods to express their opinion regarding various attributes and product summary. Further, the clients may view reviews, recommending information and sentiments regarding products and their specific attributes. These websites provide the product's summary with its features to their clients in graphical formats. Some eminent e- commerce sites such as amazon.com deliver the information based on the feedback from users with rating information and editors [19]. Sentiment Analysis Approach analyzes these large numbers of reviews and provide help to websites by converting unhappy users into supporters.

- **Voice of the Market** (VOM): This application helps determining the opinions of users regarding goods or services of competitive companies. It helps to gain benefit at competitive part and manufacturing of fresh products based on this accurate and on time information. It is essential to find this information quickly so that the marketing can be done to benefit the company.

- **Voice of the Customer (VOC)**: This method analyzes the reviews, opinions and feedback of the users, therefore very useful to get the review of every user regarding different products or services [20]. VOC contributes significantly in Customer Experience Management and provides support to find out new opportunities in the invention of new products. Additionally, some useful and ordinary requirements of products such as performance and cost are identified by retrieving the reviews of users.

- **Brand Reputation Management** (BRM): Instead of the users, it focuses on the reputation of the company, the management and their product in the market [21]. One-to- many interactions through online provides opportunities to strengthen the status of the company and their products. Companies apply sentiment analysis techniques to know about the reputation of their brands, products or amenities among online users.

- **Government**: On policy matters, the approach helps the government to take feedback of the people. Considering all these possibilities, be it through the domain of Machine Interface or Human Interaction, Sentiment Analysis plays a critical role establishing a user's view to make a clear decision. Therefore, the focus of this paper revolves around establishing possible solutions how the opinion of the customers determines the managerial process [22].

H. *SENTIMENT ANALYSIS PROCESS*

Although analyzing people's emotions is difficult, Figure 2 depicts the five tasks of data collection, text preparation, sentiment detection, sentiment classification, and output presentation as the overall approach of sentiment analysis.

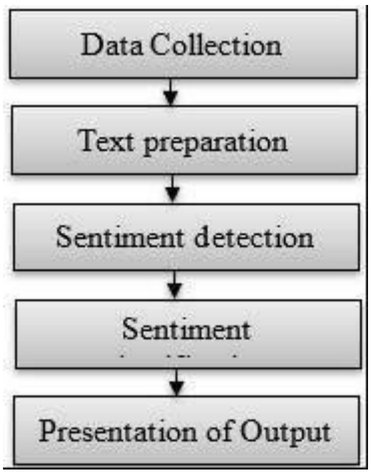

Figure 2. Sentiment Analysis Procedure

All these steps are explained below:

- **Data Collection**: The first task is to gather data from the users generated from blogs, forums and social networks which are unsystematic [23]. This includes vocabularies, slangs, context of writing etc. expressed in different ways, and not possible to analyze manually, hence the Text Analytics and NLP are applied to extract the feature and the classification

- **Text Preparation**: This effort focuses on cleaning up the extracted data before analysis. This process identifies and eliminates the non-textual and redundant contents [24].
- **Sentiment Detection**: The objective of this work is to analyze the reviews and opinions' extracted sentences. This process retains the sentences with subjective expressions but discards the sentences with objective information.
- **Sentiment Classification**: This phase entails dividing subjective statements into positive, negative, good, and bad; like, dislike, etc. However, there are a number of points on which classification depends [25].
- **Presentation of Output**: In this, converting amorphous text into meaningful information are done. After analysis, the text outcomes are exhibited on graphs. Additionally, by showing time in graph style, time may be evaluated and a sentiment time line with the chosen value is created with time.

I. MOTIVATION AND OBJECTIVES BEHIND PERFORMANCE ENHANCEMENT IN SENTIMENT ANALYSIS

Sentiments are analyzed using the polarity-based techniques such as ML and Deep Learning. Though both the techniques are efficient, it is established that if the ML techniques are applied, the SA model ought to have high accuracy for the training. It is because if the data is not clean, it can reduce the accuracy for the model training, therefore, the present study is undertaken to achieve high accuracy in sentiment analysis for which the model needs clean data.

Using deep learning algorithms, SA has been conducted with better results in recent years, and not only on text level data; attempts have also been made to extract sentiments from picture and video (action) data. Van, Tahi, and Nghiem (2017) in their paper "Combining Convolution and Recursive Neural Networks for Sentiment Analysis" state that a CNN-RNN hybrid beat all other deep neural networks. But first, let's discuss what machine learning is and why it's so effective at solving issues with large amounts of data before we get into the specific algorithms we utilized to find a solution.

J. STANDARD SENTIMENT ANALYSIS PROCESS FLOW

- Start
- Data Extraction
- Pre-Processing to reduce Objectivity
- Feature Selection Feature Weighing
- Method Classifier

II. RELATED WORK

Pang et al. explored the challenge of categorizing papers not only by subject but also by emotion in their study. Movie review has been utilized by author as data on which comparison sampling would be done using Naïve Bayes, Maximum Entropy and Support Vector Machine classifiers accordingly. Although subjects were generally discernible from keywords alone, emotion could be presented in a more straightforward manner, making for a more difficult task than standard topic-based categorization. Example: "How could anyone sit through this movie?" is completely free of negative connotations. Therefore, comprehending sentiment need more knowledge than the typical topic-based categorization. In a reiteration of previous findings (Pang, Lee, & S, 2002), support vector machine was shown to be the most effective method. To determine the overall polarity of the text for sentiment classification, Pang and Lee have created a subjectivity classifier. The classifier has boosted accuracy and provided a succinct summary of the content's sentiment analysis. A large number of features made the applications of senti- ment analysis infeasible during research work. **Tim O Kfee et.al.** gave feature selection and weighting methods by in- volving Naïve Bayes and Support Vector Machine (O'Keefe & Koprinska, 2009). Movie review sentiment was figured out using 3 feature selections, namely, 1) Categorical Proportional Difference (PD) which tells us the equality among two num- bers; 2) Sentiword Net Subjective Scores (SWN-SS) have given the sentiments of subjective terms and 3) Sentiword Net Proportional Difference (SWN-PD) has given a stable and meaningful sentiment. There have been 6 feature weighting methods, (i) Unigram feature to consider unique word as a feature, (ii) Feature Frequency to define every feature value of unigram, (iii) Sentiword Net Word Score Groups (SWN-SG) for feature presence in unigram, (iv) Sentiword Net Polarity Groups (SWN-PG), (v) Sentiword Net Polarity Sums (SWN- PS), and (vi) Term frequency. Of the various methods of feature selection, the Proportional difference (PD) and SWN- PD have achieved higher accuracy of 87.15% using Machine Learning.

In the research article by **Rudy Parabowo**, hybrid classi- fication approach has been discussed for sentiment analysis dealing with both Machine Learning techniques and Rule Based Classification. Such approaches can be used for the analysis of movie reviews, product reviews and MySpace comments; classification of the given dataset has been done through micro averaging and macro averaging to determine F1. F1 is a measure that takes precision and accuracy into account. This paper has focused on automatic document classification having set of classes for multiple documents to define a particular document belongs to which class set. Machine Learning involving two set of documents: training set for differentiating document characteristics and testing set for validating performance. Micro and macro averaging created confusion tables and thus average the input parameters to calculate recall and precision. Rule Based Classification involved antecedent and consequent providing optimization by replacing proper noun by '?' Or '#' to form pattern while Machine Learning increases accuracy by assigning weight to each feature and hence represents the document by feature set. Thus, optimum weight and other parameters have defined high level of effectiveness on unseen sample data (Prabowo & Thelwall, 2009). For the online product review a Lexicon Enhanced method for sentiment classification has been defined by **Yang Dang et.al.** The method allows internet to be user centric, such that people can exchange information through online social community like web forums and discussion boards. The re- search identifies and analyzes text containing opinion and emotions based on subjectivity or objectivity of data. The approach has been classified using supervised machine learning or semantic orientation approach having no prior training. Lexical Enhanced approach combines the efficiency of both approaches and generated set of sentiment words based on sentiment lexicon as new feature dimension and combined their sentiment feature with content free and content specific machine learning approach. The semantic orientation approach allows classification based on (i) Dictionary based approach or (ii) Corpus based approach. Dictionary based technique based on antonyms, synonyms and hierarchies in word net. Corpus based technique depending on statistical information has not been as efficient as dictionary-based technique (Dang, Zhang, & Chen, 2010).

The efficiency of Back propagation algorithm for text categorization is being discussed in the research work of **Ramasundaram and Victor** allowing the text categorization to extract the meaningful information. Back propagation neural network has been useful for complex pattern recognition and performs non-trivial mapping function pro- viding n-dimensional feature vector representation of object. To improve the accuracy of classifier the raw features have been converted to improved features using feature selection and extraction. Three fundamental steps proposed for text categorization involve feature set identification, dimensionality reduction and three-layered feed forward Back propagation network. Feature identification techniques defined like word extraction, stop word removal, stemming, giving weight to terms. To improve the scalability, dimensionality reduction has been done and hence neural network classifier has adopted binary classification to input feature space. Two phases of Back propagation algorithm have been defined; first phase in- volved forward signal propagation and second phase included error feedback to input units. Thus, algorithm as defined by the authors provides minimum optimization and simulated annealing for text categorization problem (Ramasundaram & Victor, 2010).

The sentiment classification for a given text depended upon i-model given by **Zhu Jian et.al** to define if the opinion is negative or positive. The usage of individual model, also known as the i-model is based on ANN for textual sentiment analysis. The i-model is a supervised model which includes prior knowledge, features and feature weight. The i-model gave precise predictions of dataset polarity and accurate opinions and edits the incorrect ones. The i-model provides analysis of sentiment polarity for movie reviews dataset and analyzed reviews corpus into Thumbs up (positive) and Thumbs down (negative). Contrary opinions were focused in the research (Zhu, Xu, & Wang, 2010).

To deal with feature selection for opinion classification by **Ahmed Abbasi**, different feature representation techniques have been used like bag of words (BOW), n-grams etc. A sentence or a phrase is considered as a multi-set of words which neither maintains the order of the words nor the grammar. BOW is used for information retrieval (IR) and NLP (Manning, Raghavan, & Schutze, Introduction to Information Retrieval, 2008). Feature selection methods exploited syntactic properties of text and supported relevant semantic information. Feature sub Sumption hierarchies (FSH) and excellent Intelligent Feature Selection (IFS) approach have removed irrelevant higher order n-grams. IFS have refined large input feature space through part of speech, word n-grams, and character n-grams. Information Extraction Pattern (IEP) is defined for three datasets of online prone product reviews-digital camera, automobile and movie reviews using 5-fold cross validation (Abbasi, 2010).

As given by researchers **Georgios Paltoglou & Mike Thelwall,** TF-IDF makes Sentiment Analysis more accurate on a wide variety of selected datasets. An unorderly collection of words in a document were modelled by adopting the bag of words (BOW) representation. The TF- IDF approach improved the classifier's performance and no annotation or external sources were required (Paltoglou & Thelwall, 2010).

Social Networking and Micro blogging sites are storage of human opinions. The researchers **Joshi, A., Balamurali, A. R., Bhattacharyya, P., Mohanty, R.** have developed a system C- Feel-It. It takes opinion on the basis of contents in tweets from micro blogging site twitter. This web-based system categorizes tweets as positive or negative (subjective) or objective and generates an aggregated sentiment score giving live snapshot of sentiment of user about topic. It is a rule-based system developed to classify tweets. A weighted majority voting principle has been used to predict the sen- timent of tweets. Finally, C-Feel-It has divided tweet state- ments into three parts-(1) Tweet Fetcher (2) Tweet sentiment predictor and (3) tweet sentiment collaborator (Joshi, A R, Bhattacharyya, & Mohanty, 2011).

III. MATERIALS AND METHODS

Opinionated terms that are distributed across multiple sentences and identifying such terms in a sentence is important for opinionated content identification. The Word Embedding Self-Attention LSTM couldn't handle multiple sentences. To overcome these limitations and to improve the performance of the model on the dataset having multiple sentences, a Universal Sentence Encoder (USE) integrated with a Self-Attention based LSTM neural network (SE-SALSTM) is introduced.

3.1 Sentiment Embedding-Self Attention LSTM

SE-SALSTM correlates the terms that are distributed in different parts of the review and assigns high weights to those terms for better prediction. The proposed SE-SALSTM architecture is shown in Figure 5.1. This architecture is comprised of Universal Sentence Encoder, Long Short-Term Memory and Self-Attention. The functionality of each layer is examined in the below section.

A. UNIVERSAL EMBEDDING-SELF ATTENTION LSTM

USE assigns weights to a word based on the significance of the word across multiple sentences in the review. It uses a Deep Averaging Neural Network (DAN) for sentence embedding.

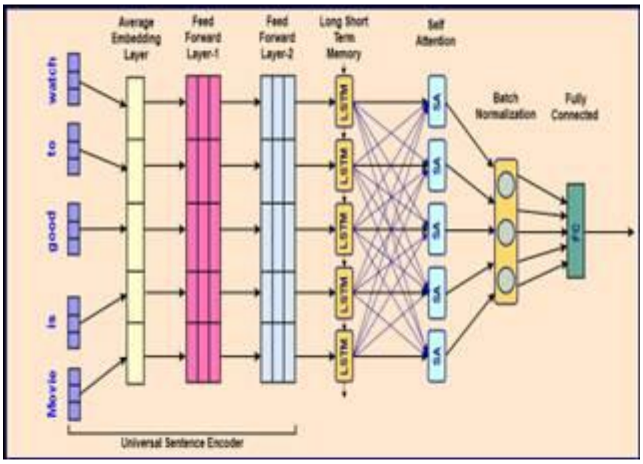

Figure 3. Proposed architecture of Sentence Embedding and Self Attention based LSTM (SE-SALSTM).

This network is preferred for embedding because it avoids duplicate phrases in a sentence and ignores the order of words. DAN is composed of the BOW layer, Average Embedding layer and two or more Feed Forwarding layers. The first layer converts the tweets into a word vector of 0s and 1s. This vector is fed into an average embedding layer where it finds the mean of all word vectors. The mean vector is passed through a sequence of Feed-Forward layers. Back propagation is applied to reduce the error rate and to fine-tune the model. It increases the generalization of the neural network. The output of USE is fed into LSTM.

The offensive words *"racists"*, *"useless"* and *"conservatives"* are appearing in different sentences of review four, USE assigns weights based on the importance of those terms in the review. The feed forward neural network extracts the complex features that exist between the terms in a review. The feature vectors from USE are fed into the LSTM to find the distance relationship between the features.

B. ***LONG SHORT TERM MEMORY***

Convolutional Neural Network (CNN) can capture only the local feature in a sentence. To capture global features, a deep convolutional neural network is required. The features located in distant parts of a sentence that is influencing the review sentiment are not captured by CNN. To overcome this limitation, LSTM is used to extract the features from arbitrary varying input. The features that are situated in distinct sections of a sentence are caught and better represented by LSTM. LSTM connects these portions and helps in predicting the opinionated terms better. It has special Gates that serves to remember the important terms and forget the irrelevant terms. The word *"politicans"* are related to *"racists"* and *"useless"* in the first sentence, while *"voters"* is related to *"conservatives"* in the second sentence by the LSTM to gain improvement in further stages.

C. ***SELF-ATTENTION***

LSTM network associates the distinct sections of the sentence, but they don't differentiate the parts that are relevant or irrelevant. This mechanism connects different relevant parts of the tweet and assigns a high score to those parts that contribute to the prediction of the tweet. This layer assigns a high score to the noun, adverb and adjective terms of the offensive tweet. The term *"racists"*, *"useless"* and *"conservatives"* represents noun, adjective and noun parts of speech. They assigned a high score to this term and enable prediction to be offensive. The attention visualization for the given offensive tweets is shown in Figure 5.2. In this visualization, offensive words like "*racists*", "*useless*" and "*conservatives*" are getting more attention in SE-SALSTM than in SE-LSTM.

IV. EXPERIMENT RESULTS AND DISCUSSIONS

The significance terms are distributed across multiple sentences in different parts of the review. These reviews are not weighted properly with the Word Embedding Self-Attention LSTM. To overcome these limitations and to address the significance terms a Sentence Embedding Self-Attention LSTM (SE-SALSTM) is proposed. Three experiments are conducted to find the importance of the proposed architecture. The experiments included in the section viz., Sentence Embedding-Bidirectional LSTM (SE-BLSTM), Sentence Embedding-Convolution Bidirectional Recurrent Neural Network (SE-CBRNN) and Sentence Embedding-Self Attention LSTM (SE-SALSTM). The experiments are conducted with OLID, IMDB and Polarity datasets. The explanation of the datasets is given in the section 3.0.1. We have used a Universal Sentence Encoder (USE) for sentence embedding. A deep averaging network is employed in USE that produces an output vector of 512 dimensions.

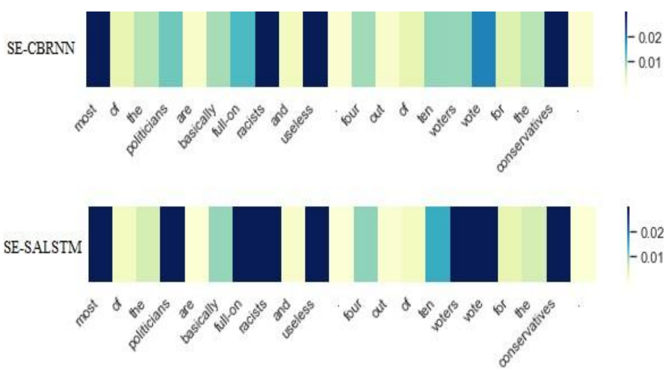

Figure 4. Proposed architecture of Sentence Embedding and Self Attention based LSTM (SE-SALSTM).

The embedding layer output is fed as input to the LSTM layer with 512 cells.

4.1 Sentiment Embedding-Self Attention LSTM

Word embedding neural networks cannot find the features that are distributed across multiple sentences. To break the distance between any two words in a tweet, Universal Sentence Encoder (USE) is used in this architecture. A transfer learning-based approach is applied in this experiment. The output of Universal Sentence Encoder is fed into

Bidirectional Long Short Term Memory (SE-BLSTM). The unidirectional model processes the input from the forward direction only, whereas the bidirectional model processes in both the directions. The bidirectional model better understands the contextual information. An offensive tweet is taken from OLID dataset,

Review 6: *"most of the politicans are basically full-on racists and useless. four out of ten voters vote for the conservatives."*

Above tweet conveys an offensive message about the politicians. The Universal Sentence Encoder selects the key phrases from different sentences of the review like *"most politicans"*, *"basically full"*, *"racists useless"* and *"vote conservatives"*. As the BLSTM approach scans the tweet in both directions from left to right and right to left, it better understands the context of the tweet as offensive. They can better understand the opinionated words than the unidirectional model. In our first experiment, we have analyzed the performance of SE-BLSTM. This network has obtained an F_1 value of 79.15%, 83.93% and 81.22% for OLID, IMDB and Polarity datasets respectively. The summary of the experiment in terms of performance is listed in Table.1.

Table 1. Performance of SE-BLSTM on IMDB, Polarity and OLID dataset.

	IMDB	*Polarity*	*OLID*
Accuracy	83.12	81.65	80.48
Precision	89.95	80.98	75.6
Recall	78.66	81.47	83.05
F_1 Score	83.93	81.22	79.15

Table 2. Performance of SE-CBRNN on IMDB, Polarity and OLID dataset.

	IMDB	*Polarity*	*OLID*
Accuracy	88.98	83.41	81.65
Precision	89.95	81.58	77.99
Recall	87.85	84.09	83.48
F_1Score	88.89	82.82	80.64

4.2 Experiment-2 Sentence Embedding and Hybrid CBRNN

In opinionated content identification, automatic extraction of word-level features is found to be less important than sentence-level features. To combine the sentence- level feature extraction and context-based classification, a hybrid model is integrated with sentence embedding. This hybrid model fuses the properties of CNN and BGRU to provide a unique model for the classification of the reviews. BGRU is a variant of Bidirectional RNN that has a smaller number of gates than BLSTM and it does not have internal memory. The sentence-level features are extracted with a Universal Sentence Encoder on the top. The CBRNN at the bottom relates the semantic information distributed in different parts of the sentences to provide classification.

At the top Universal Sentence Encoder selects the key phrases from different sentences of the review like *"most politicans"*, *"basically full"*, *"racists useless"* and *"vote conservatives"*. This key feature is given to the next CBRNN layer. The BGRU runs either direction of the review, uses the convolution layer extracted chronological feature to predict the sentence as offensive. For SE-CBRNN model, we achieved an F_1 score of SE-CBRNN with the value of 80.64%, 88.89% and 82.82% for the OLID, IMDB and Polarity datasets, respectively and it is shown in Table 2.

4.3 Experiment-3 Sentence Embedding and Self-Attention Sequence Model

To assign high weights to the pertinent parts of the review and increase the overall weight of a review, we have integrated sentence embedding and LSTM with self-attention.

Table 3. Performance of SE-SALSTM on IMDB, Polarity and OLID dataset.

	IMDB	*Polarity*	*OLID*
Accuracy	96.91	82.65	80.3
Precision	81.25	78.95	90.91
Recall	98.67	84.62	74.51
F_1Score	89.12	81.68	81.9

In the tweet, the phrases like *"racists useless"* and *"conservatives"* located in different sentences are related by USE while calculating weights of those phrases. LSTM relates the different parts of the longer sequence to compute the strength of a word. In the above example, the phrases like *"racists useless"* and *"conservatives"* represents noun-adjective and noun relationship respectively. When the self-attention evaluates the weight of *"conservatives"* phrase, it picks up the weight of those phrases that are arbitrary located at a distance and uses them in the evaluation. By using this evaluation method, it has obtained an F_1 Score of 81.9%, 89.12%, 81.68% on OLID, IMDB and Polarity datasets respectively and it is shown in Table 5.3. Instead of using a Word Embedding Self-Attention LSTM, we have used a Sentence Embedding Self-Attention LSTM to find the relative terms across the multiple sentences and to assign high weights to those terms for better opinionated content identification and categorization.

A. COMPARISON BASED ON METRICS

BLSTM is a type of Bidirectional Recurrent Neural Network capable of processing the sequence input effectively than CNN. It can capture global features rather the local features. i.e., terms that are separated by a distance. Based on these features SE-BLSTM has obtained an F_1 score of 79.15%, 83.93% and 81.22% on OLID, IMDB and Polarity datasets respectively. To improve the performance in identifying the opinionated content based on context, we have integrated Universal Sentence Encoder and Convolution Bidirectional Recurrent Neural Network (SE-CBRNN). This SE-CBENN is capable of extracting the key set of phrases across multiple sentences in a review and scans the review in both directions to provide a context-based classification.

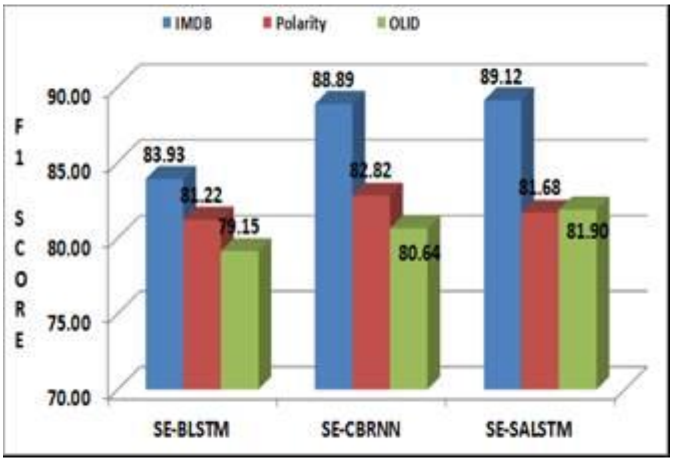

Figure 5. Proposed comparison of proposed architecture with base models based on F_1 score

Table 4. Performance comparison of proposed architecture with baseline models based on F_1 score**.**

	SE-BLSTM	*SE-CBRNN*	*SE-SALSTM*
IMDB	83.93	88.92	89.12
Polarity	81.22	82.82	81.68
OLID	79.15	80.64	81.90

Table 4 and Figure 5 shows the F_1 score of SE-CBRNN with the value of 80.64%, 88.89% and 82.82% on the OLID, IMDB and Polarity datasets respectively. To improve the performance in identifying the opinionated content, we have integrated Universal Sentence Encoder and Self-Attention with LSTM (SE-SALSTM). This SE-SALSTM is capable of finding the key set of phrases across multiple sentences in a review and assigns high weight to the pertinent parts of the review. This proposed neural network finds the relative terms across the sentences and assigns high weights to those terms that influence the sentiment of a tweet. This proposed SE-SALSTM achieved an increase in F_1 score of 2.75% and 1.25% than SE-BLSTM and SE-CBRNN on the OLID dataset.

B. COMPARISON BASED ON EXECUTION TIME

The training and testing time of the proposed model with the baseline model is shown in the Table 5 for the OLID, IMDB and Polarity datasets. The time taken to fit the model with data as training time, while predicting the data as positive or negative as testing time. SE-BLSTM is slower in processing the opinionated content because it builds the relationship between hidden vectors for each timestep. This neural network takes a training time of 25.34 minutes on the OLID dataset. In SE-CBRNN model, Universal Sentence Encoder takes time to construct an embedding matrix for the phrase using feed forward neural network and CBRNN model in bottom takes time to find the chronological features. The SE-CBRNN takes a training time of 26.04 minutes on the OLID dataset. To improve the performance of neural networks in identifying opinionated content, we have integrated Self-Attention with sentence embedding - Universal Sentence Encoder. This Universal Sentence Encoder has a two-layer feed-forward neural network to find the complex relationship between the data and an attention mechanism adds weights to the model, which indirectly increases the training time of the data. So, it takes a training time of 27.53 minutes on the OLID dataset.

Table 5. Performance comparison of proposed architecture with baseline models based on training and testing time.

		SE-BLSTM	*SE-CBRNN*	*SE-SALSTM*
OLID Dataset	Train Time (mins)	25.34	26.04	27.53
	Test Time (mins)	7.25	7.56	8.31
IMDB Dataset	Train Time (mins)	291.32	294.21	312.54
	Test Time (mins)	59.34	58.52	62.46
Polarity Dataset	Train Time (mins)	22.51	23.34	24.32
	Test Time (mins)	6.52	6.24	7.25

V. CONCLUSION

In conclusion, tweet sentiment analysis with Python and machine learning is a powerful tool that can help businesses and individuals gain valuable insights into public opinion. By using machine learning algorithms to analyze large volumes of tweets, we can quickly determine the overall sentiment towards a particular topic, product, or brand. With the help of Python's data analysis and machine learning libraries, we can easily preprocess the tweet data, extract relevant features, and train a machine learning model to accurately classify tweets based on their sentiment scores. Once the model is trained, we can use it to make predictions on new, unseen tweets in real-time. Overall, tweet sentiment analysis is a fascinating and rapidly evolving field that has numerous applications in business, politics, and social media. With the right tools and techniques, we can gain valuable insights into public opinion and use this information to make data-driven decisions.

REFERENCES

[1] M. D. Devika, C. Sunitha, and A. Ganesh, "Sentiment Analysis: A Com- parative Study on Different Approaches," *Procedia Computer Science*, vol. 87, pp. 44–49, 2016.

[2] A. Vishal, S. S. Kharde, and Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques"," *International Journal of Computer Applications*, vol. 139, pp. 11–21, 2016.

[3] A. Go, R. Bhayani, and L. Huang, *Twitter Sentiment Classification using Distant Supervision*, Stanford, 2009, vol. 46.

[4] Y. Kawai, T. Kumamoto, and K. Tanaka, "Fair news reader: Recom- mending news articles with different sentiments based on user pref- erence," *Proceedings of Knowledge-Based Intelligent Information and Engineering Systems (KES), number 4692 in Lecture Notes in Computer Science*, pp. 612–622, 2007.

[5] A. Kennedy and D. Inkpen, "Sentiment Classification of Movie Reviews Using Contextual Valence Shifters," *Computational Intelligence*, vol. 22, pp. 110–125, 2006.

[6] J. Spencer and G. Uchyigit, "Sentiment or: Sentiment Analysis of Twitter Data," *James Spencer and Gulden Uchyigit*, vol. 88, pp. 1–11, 2016.

[7] "Doaa Mohey El-Din Mohamed Hussein "A Survey on Sentiment Analysis Challenges"," *Journal of King Saud University - Engineering Sciences*, vol. 50, pp. 1–9, 2016.

[8] P. Shahnawaz and Astya, "Sentiment Analysis: Approaches and Open Issues," *International Conference on Computing, Communication and Automation*, vol. 9, pp. 1–5, 2017.

[9] L. Hu, "Mining and Summarizing Customer Reviews," in *International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 168– 177.

[10] A. M. Popescu and O. Etzioni, "Extracting Product Features and Opin- ions from Reviews," in *Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 2005, pp. 339–346.

[11] P. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, 2008, pp. 417–424.

[12] M. Annett and G. Kondrak, "A comparison of sentiment analysis techniques: Polarizing movie Blogs," *Canadian Conference on AI*, pp. 25–35, 2008.

[13] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[14] Abbasi, A, Hsinchun C, and Arab S., "Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums" ACM *transactions on information systems (TOIS)* 26, no. 3, 2008, PP.1-34.

[15] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Systems with Applications*, vol. 36, pp. 6527–6535, 2009.

[16] B. Liu, "Handbook of Natural Language Processing," *Chapter Sentiment Analysis and Subjectivity*, pp. 1–38, 2010.

[17] Liu, B., 2010. "Sentiment analysis: A multi-faceted problem. *IEEE intelligent systems*," vol. 25, pp. 76–80, 2010.

[18] J. Fang and B. Chen, "Incorporating Lexicon Knowledge into SVM Learning to Improve Sentiment Classification," *Proceedings of the Work- shop on Sentiment Analysis where AI meets Psychology (SAAIP)*, pp. 94–100, 2011.

[19] V. Rout and A. D. Londhe, "Survey on Opinion Mining and Summa- rization of User Reviews on Web," *International Journal of Computer Science and Information Technologies*, vol. 5, pp. 1026–1030, 2014.

[20] A. Pelleg and Moore, "X-means: Extending K-means with Efficient Estimation of the Number of Clusters," in *Proc. of the 17th Int. Conference on Machine Learning*, 2000, pp. 727–734.

[21] B. Pang, L. Lee, and S. Vaithyanathan.

[22] E. Riloff and J. Wiebs, "Learning Extraction Patterns for Subjective Expressions," in *Natural Language Processing*, 2003, pp. 105–112.

[23] T. Wilson and J. Wiebe, "Annotating opinions in the world Press," in *4th SIG dial Workshop on Discourse and Dialogue*, 2003, pp. 13–22.

[24] B. Pang and L. Lee, "A sentimental education: Sentimental analysis using Subjectivity Summarization based on Minimum cuts," in *Pro- ceeding of the 42nd Annual Meeting on Association for Computational Linguistics*, 2004, pp. 271–278.

[25] A. Esuli and F. Sebastiani, "PageRanking WordNet Synsets: An Ap- plication to Opinion Mining," in *45th Annual Meeting-Association for Computational linguistics*, vol. 45, 2007, pp. 424–431.

[26] B. Pang and L. Lee, "Opinion Mining and Sentimental Analysis," *Foundations and Trends in Information Retrieval, USA*, vol. 2, no. 2, pp. 1–135, 2008.

[27] A. Beygelzimer, J. Langford, and B. Zadrozny, pp. 3–28, 2008.

[28] S. Cai, Y. Spangler, L. Chen, and Zhang, "Leveraging Sentiment Anal- ysis for Topic Detection," *International*

Conference on Web Intelligence and Intelligent Agent Technology, pp. 265–271, 2008.

[29] A. Agrawal, F. Biadsy, and K. R. Mckeown, "Contextual Phrase- Level Polarity Analysis using Lexical Affect Scoring and Syntactic n- grams," in *Proceeding of the 12th Conference of the European Chapter*. Association for Computational Linguistics, 2009, pp. 24–32.

[30] S. Li, H. Zhang, W. Xu, G. Chen, and J. Guo, "Exploiting Combined Multi-Level Model for Document Sentiment Analysis," *International Conference on Pattern Recognition*, pp. 4141–4144, 2010.

[31] R. Lau, W. Zhang, P. Bruza, and K. Wong, "Learning Domain-specific Sentiment Lexicons for Predicting Product Sales," *IEEE International Conference on e-Business Engineering*, pp. 131–138, 2011.

[32] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, 2011.

[33] M. Karqmibekr and A. A. Ghorbani, *Sentiment Analysis of a Social Issues*, USA, 2012.

[34] M. Eirinaki, S. Pisal, and J. Singh, "Feature-based opinion mining and ranking," *Journal of Computer and System Sciences*, vol. 78, pp. 1175-1184, 2012.

[35] P. H. Shahana and B. Omman, "Evaluation of Features on Sentimental Analysis," *International Conference on Information and Communication Technologies*, pp. 1585–1592, 2015

[36] U. Farooq, T. Prasad, A. Nongaillard, Y. Ouzrout, and M. Qadir, "A Word Sense Disambiguation Method for Feature Level Sentiment Analysis," *International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, pp. 1–8, 2015

[37] M. Jayraj, Desai, R. Swapnil, and Andhariya, "Sentiment analysis approach to adapt a shallow parsing-based sentiment lexicon," *2015 International Conference on Innovations in Information, Embedded and Communication Systems*.

[38] S. Ahmed and A. Danti, "Effective Sentimental Analysis and Opinion Mining of Web Reviews Using Rule Based Classifiers," *Advances in Intelligent Systems and Computing*, vol. 410, pp. 171–179, 2016.

[39] M. Karqmibekr and A. A. Ghorbani, *Sentiment Analysis of a Social Issues*, USA, 2012.

[40] M. Eirinaki, S. Pisal, and J. Singh, "Feature-based opinion mining and ranking," *Journal of Computer and System Sciences*, vol. 78, pp. 1175–1184, 2012.

[41] E. Haddia, X. Liua, and Y. Shib, "The Role of Text Pre-processing in Sentiment Analysis," *Information Technology and Quantitative Management*, vol. 17, pp. 26–32, 2013

[42] K. Ghag and K. Shah, "Comparative Analysis of the Techniques for Sentiment Analysis," *International Conference on Advances in Technology and Engineering (ICATE)*, pp. 1–7, 2013.