E-COMMERCE PRODUCT BASED REVIEW AND NEW STRATEGY FOR SENTIMENT ANALYSIS

Dr Saurabh Gupta

G H Raisoni Institute of Engineering & Business Management, Jalgaon
saurabh.gupta@raisoni.net

Abstract: Customers nowadays routinely share their thoughts on social media about any product, brand, or experience. Analysts collect and analyse these reviews in order to learn more about the product. The beauty of social media is that it connects people from all walks of life. As a result, the analysts gathered feedback from various social media and platform sources for practically everything. Sentiment Analysis is used to forecast outcomes to obtain relevant information, such as predicting a movie's box office success and rating new products. In today's competitive world, this form of prediction helps customers look to purchase goods or services. The goal of this article is to review e-commerce websites in a text format with some special characters and symbols (emojis). In terms of context, emotion, and prior experience, every words has some meaning. These factors influence some of the features of text data that work for prediction. This paper aims to bring together previous research on text analysis with emotion-based analysis. There is also a discussion of outstanding issues and limitations of document-based sentiment analysis. This study came to a revolutionary multi-class categorization approach as its conclusion. A ternary classification is classes divided into positive, negative, and neutral based on product-based text and emoji evaluations on the Twitter social media platform.

Keywords : machine learning ,Sentiment Analysis, reviews, , Real time, e-commerce**.**

INTRODUCTION

Sentiment Analysis is a natural language processing (NLP) technique for determining whether an input is favourable, negative, or neutral. Sentiment Analysis is a type of textual data analysis used by organizations to track things like brand and product sentiment, also customer input, and needs.

Attitude, emotion, and experience are the three main components of sentiment analysis. Customer input on a given product is the initial step in sentiment analysis. Various users or customers have differing perspectives on the same product. Also, to gain a complete view of any product, get feedback from various channels. For two reasons, this study employs Twitter as a social media platform to gather these thoughts: one, it is popular these days, and second, it is an exact and convenient analysis because it limits the review to 140 characters.

Customers' feedback is beneficial to both other customers and sellers. Customers may acquire information about goods in the available quantity and quality by knowing the details about any product. Sellers benefit from this form of feedback since it allows them to evaluate what was good and what was bad about their product, allowing them to maintain inventory and quality for a successful business. The reviews are written in a free-form text format.This data contains some broad terms that categorize our viewpoints into several groups. In multi-class sentiment categorization, opinions are divided into strong positive, strong negative, positive, negative, and neutral categories [9]. Strong positive opinion as sentiment context such as "good, best, super, superb, and so on". Similarly, terms like awful, worse, not good, dissatisfied, filthy, and others reflect a product's unfavourable impact, which classify as a strong negative opinion. The biggest issue emerges when consumers write two statements joined by "and," "or," and "but" connectors in conjunction form. One component might be favourable for one aspect, while the other might be negative for another.When it comes to such reviews, it's difficult to say if they're positive or negative. The review do not explicitly state an opinion regarded as impartial [3].

Customers occasionally utilize unusual symbols when writing reviews, and these symbols have distinct meanings. These symbols will be transformed into text for processing, after which they will be classified into specific opinion classes, resulting in accurate predictions [4].

Overview of Machine Learning

In emotional analysis, machine learning approaches initially used. Machine learning shown in Fig-01 is a set of algorithms used to analyze sentiment in social media text data. In the following paragraphs, we'll go through some of the most popular machine learning algorithms in this area.

A. **Supervised Learning** – In this strategy, the dataset used to train the machine using previously labeled data, resulting in a more accurate mapping of new data to input and output. The main purpose of this to train machines to learn "how to map the input to output."
Also, to gain a complete view of any product, get feedback from various channels. For two reasons, this study employs Twitter as a social media platform to gather these thoughts: one, it is popular these days, and second, it exacts and convenient analysis because it limits the review to 140 characters.

Various kinds of supervised leaning:

- **Regression:** The problem is determining the amount of feature participation, utilized regression. Regression used to solve this type of problem. The key benefit of this method is that it allows you to find the continuous value of a feature. The algorithm should avoid overfitting by using a probabilistic interpretation of this value.

Based on the example above, if the machine can identify a good route for a specific period, it may use time as the primary factor in finding a suitable solution. That is the path that takes the least amount of time to travel (it takes the least amount of time) is subject to regression.

  - Linear regression
  - Logistic regression
  - Polynomial regression
  - Stepwise regression
  - Lasso regression
  - ElasticNet regression

- **Classification**: Classification is the process of dividing or grouping the output data into categories. That is, outputs of the same type belong to the same class.

  UGC CARE Listed (Group -I) Journal Volume 11, Iss 3, 2022

If the system is able to find a route based on traffic in the above scenario. Only two scenarios are possible: less traffic and more traffic. This issue divided into two categories: less traffic and more traffic. That is a type of problem that has a classification. The outputs were divided into categories/ classes.

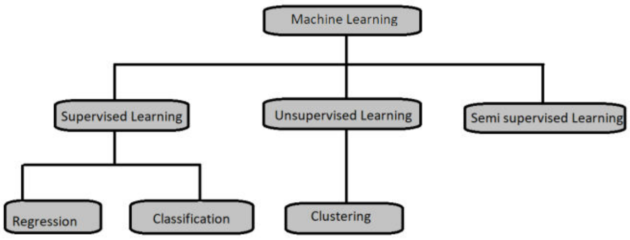

Fig. 1.Classifiers in Machine Learning

B. Unsupervised Learning – This module trained data by using unlabelled data. In this case, the machine must learn for itself. Unsupervised learning isn't classified or labeled; instead, learn from personal experience. It was a tough assignment [9].

Consider the case of a baby and his family dog. Baby is aware that a dog has two eyes, two ears, and four legs on which it walks. Assume that a new pet (a cat) has arrived in the house. It looks like a dog to a baby since it has the same features learned previously. It falls into the dog category, according to his knowledge.

Unsupervised learning aids in the discovery of valuable features for categorizing test data. This method uses unlabeled data from a self-learning model to label it. The machine will look for a pattern in a given data set to solve the problem.

Unsupervised learning comes in a variety of forms:

- **Clustering:** This method works by grouping unlabeled data. That is, data of comparable sorts are a group (cluster). It clearly states that all data in a group is more comparable properties and are distinct from other groups. This technique utilizes in image processing in the biological study for patterns, text data structured into a topic hierarchy based on content, spam filtering, criminal activities, and sports. Among the most well-known clusters are:
  - K-Mean clustering
  - Mean- shift clustering
  - K-NN(k nearest neighbor)
  - Hierarchical clustering
  - Probabilistic clustering
- **Association Rules:** This technique builds a rule-based system. This method uses to determine the relationship between variables in a given dataset. This strategy often utilized all of the relationships in the dataset. The following are some examples of basic association techniques::
  - Apriori algorithm
  - FP-Growth Algorithm
  - Dimensionality reduction

C. **Semi-supervised Learning** – Unlabeled data is learned through labeled data for testing in this form of learning. The main goal of a semi-supervised strategy is to label a large amount of unlabeled data using small amounts of labeled data [9].

Consider the following scenario: a child has 40 vehicles (toys), but his elders only recognize four of them as cars. Now the child will be able to find the remaining 36 toys labeled as cars on his own.

The supervised and unsupervised parts of a semi-supervised approach divides by a thin line. To improve the accuracy of this boundary, it should include some learning dataset properties:

- Size of unlabeled portion: the unlabeled portion must be large enough, which implies label data should account for 4-5 percent of the whole data set. When the size of the label data increases, it falls under the category of supervised learning.
- Unlabeled data is used to generate an output based on the similarity of the labelled data. All data of the same type must be in close proximity to one another.
- Labeling complexity and simplicity: when the hidden meaning of labeled data is complicated, our situation becomes more complex than it was previously.

Inductive and transductive learning: inductive learning entails developing a generic categorization rule and applying it to test examples. Transduction isn't based on a generic rule; rather, it creates reasoning for specific training instances and applies it to specific training test scenarios.

In semi-supervised learning, algorithms utilize:

- Self training
- Generative
- Low-density Separation
- Practical application

Sentiment analysis was done using supervised and semi-supervised learning. Binary and multiclass learning techniques are the two types of learning techniques. The categorization algorithm for three types of feelings is proposed in this research.

This paper's structure is broken into six sections. The first section deals with the introduction and definition of the problem. The second portion looks at existing reviews research also sentiment analysis from the extensive poll. Design and explain the features of existing modules as well as the new module in Section 3. Section 4 summarises the previous, current, and most recent implementation methods related to the planned sentiment analysis study. Section 5 offers a comparison with other studies, followed by section 6's conclusion.

RELATED WORK

The crucial and relevant issues are discussed in this section of the paper's literature. In text analytics, the two most common approaches are machine learning and natural language processing. Several references in this work describe the approaches of supervised learning, unsupervised learning, and semi-supervised learning.

The performance of machine learning algorithms determined by several factors, one of which is data representation. Different interpretations will interact more or less, and different meaning will cover the facts, depending on how well you comprehend the data.

Domain-specific information utilized to build depictions, learn with generalization, and search for AI, which motivate the development of more efficient algorithms.

Sentence Based Sentiment Analysis (SBSA), Document-Based Sentiment Analysis (DBSA), Aspect Based Sentiment Analysis (ABSA), and Comparative Sentence Analysis (CSA) are the four types of sentiment analysis.

B. Seetharamulu and colleagues developed a deep learning-based methodology for classifying positive/ negative customer feedbackB. Seetharamulu et al. developed a classifier using supervised machine learning techniques. To determine the sentiment orientation of each text, it uses deep learning-based sentiment analysis (DL-SA)[1].

The deep neural network method's purpose learns to analyze data at a high level while avoiding horrible tasks. The availability of large-scale training data is critical to the model's effectiveness. This model is based on deep learning and recognizes emotions (DLER). The CNN technique train data on prior information to predict desired results for neural networks as part of the deep learning process. This model compares DL-SA to a previously published model for Weakly Supervised Definition Extraction WDE [1].

Azwa Abdul Aziz and his colleagues concentrated on opinion mining based on client feedback or remarks. According to the author, in-domain sentiments operate perfectly fine. However, model performance in real-time data sets may be affected, and SML may drop in cross-domain datasets. These models are based on Contextual Analysis (CA), which allows for the construction of hierarchical knowledge tree structures from links between words and sources (HKT). Azwa Abdul Aziz et al. developed the Tree Similarity Index (TSI) and Tree Differences Index (TDI). The main goal is to develop performance measures for SML models in real-time data, and then construct a relationship in CA that can utilised to interpret structured data knowledge [2].

SML is a text recognizer for data comparison by Azwa Abdul Aziz et al. SML uses training data to train a classifier that predicts the desired outcome. The majority of the classifier trained using n-grams, which is a contiguous sequence of n items from a particular sample text. Because the CA notion is based on text, it gets transformed to HKT. HKT employs unlabeled data to aid in comprehension of subject and data expertise. In SML, the CA technique is used to predict the success or failure rate of other models. Text normalization, this model employs two methods: first stemming and second lemmatization [2].

To obtain a stem, stemming removes affixes such as suffix, prefix, and infix from the word's order. The purpose of the lemmatization procedure is to group the various inflected forms of words it may be studied as a single text. Cross-domain replication were studied by Azwa Abdul Aziz et al. [2].

For comparative polarity analysis based on the Senti approach, Kartikkayini T and N K Srinath employed the NLTK and Datumbox tools. Kartikkayini T and N K Srinath centred their inquiry on three levels: the first document, the second sentence, and the final opinion mining classification using subjective and objective criteria. This model distinguishes between positive polarity, negative polarity, and neutral polarity. This approach begins with raw data, which is passed through an IDE (RStudio or Python) to a JSON parser, which analyses the data before feeding it to NLTK for sentiment analysis and finally converting it into a chart [3].

Overall product reviews are divided into three categories by the Senti algorithm: sentiment score, partial sentiment score, and sentence score. The average of partial sentiment and sentence scores will be used to compute the overall score [3].

The review analysis was overseen by Ronen Feldman. According to Ronen Feldmen, the sentiment is obtained from social media sites such as Facebook, message boards, blogs, forums, and tweets. Ronen Feldman specialises in sentiment analysis at the document, phrase, and aspect levels, as well as comparative sentiment analysis and sentiment lexicon acquisition. A corpus of documents input into the system is transformed to text and pre-processed using linguistic methods like stemming, tokenization, part of speech tagging, entity extraction, and link extraction, according to Ronen Feldman. Before processing, the document analysis module used linguistic resources to annotate texts. Several visualisation programmes include this annotation. [4].

K. Saranya and S. Jayanthy presented art that combined textual sentiment with emotion. Sentiment classifications based on ontologies mostly bring emotions to the text. This model accepts emotional text as input, which must be rated as either positive or negative. Emotion texts are feds into the NLP system, which analyses and annotates the text [5].

K. Saranya and S. Jayanthy analyze a dataset as movie reviews and categorize them as good, negative, or neutral. The dataset was built by classifying texts using thumbs-up and thumbs-down signals. Emotions were ascribed to semantic roles for the creation of emo words by the models, which found an affective class hierarchy in WordNet. These emo words are used to create an emotional hierarchy. For a better outcome, k. Saranya and S. Jayanthy employed a mix of SVM and Navie-Bayes algorithms. [5].

Nadijim Frechet and his colleagues are working on syllabi from the world's top 50 universities, according to the QS World University Rankings. This method used 1000 different sources and produced indexing to determine how a publication was used, which is a type of citation. As a result, it may use cutting-edge machine learning techniques to classify data into appropriate sectors. In order to understand Wordscores, the conceder also gathered text data from audio and video [6].

R and Python are compared by Nadijim Frechet et al, who conclude that Python is preferable for large-scale or prediction-based applications. They employed R tools like tidyverse, tidytext, stm, and quanteda, as well as Python packages like pandas, NumPy, Scikit-learn, NLTK, spaCy, and Gensim in their research [6].

Dr. B Valarmathi and K Sudheer are working on real-time sentiment analysis of Twitter consumer reviews. Gather evaluations from e-commerce sites and categorise them using machine learning algorithms. The key subjects of this assignment [7] are sentiment lexicons, sentiment analysis, and machine learning models.

Dr. B Valarmathi and K Sudheer downloaded live tweets using the Twitter Streaming API. Amazon has 50000 Twitter followers, e-bey has 25000, and Alibaba has 25000. This data needs be stripped of retweets, URLs, the # tag, and special characters. The NLTK libraries were used in this model for stemming, tokenization, machine learning, parsing, and semantic reasoning. More than 50 corpora and lexical resources are used in the sentiment analysis approach [7].

This model used Document Frequency (DF) to discover frequently occurring words in data using a predetermined threshold range. This model contrasts Navie Bayesian, max-entropy, and decision rule-based categorization [7].

Lijuan Huang et al. developed a sentiment model for polymerization topics (PTSM). To get sentiment data from internet reviews, the PTSM model extracts and filters data. This model proposed hidden sentiment information at the document level. This method addresses the over-fitting problem by filtering out unnecessary information from online reviews, making it straightforward to apply a machine-learning algorithm to provide accurate predictions [8].

Using a corpus method and lexicon approach, Lijuan Huang et al. were the first to develop a data dictionary. This data dictionary used manual data to identify the most positive and negative relevant traits, and then used word segmentation to separate them into two groups (WS). Pass the dictionary through a neural network (NN) and a support vector machine (SVM) before running it through PTSM to get a more accurate prediction than earlier approaches [8].

Abinash Tripathy et al. use text reviews to find product predictions in this model, which employs a mixed machine learning method. First, this model uses the SVM technique to generate output use as input for the sentiment classification procedure using an artificial neural network (ANN). This model utilised the Term frequency-inverse document frequency (TF-IDF) and CountVectorization (CV) methods to find features using the IMDb data set and the polarity dataset [9].

When the frequency of words increases in a document, the TF-IDF value increases in this model. The term frequency refers to the number of times a word appears in the text, whereas the inverse document frequency refers to the term's appearance across the document. The text document converte to sparse matrices via CountVectorization [9].

This model uses two movie reviews datasets, IMDB and polarity takes an input and applies a k-fold cross-validation method. Preprocessing is data first to remove any stop words such as numerical values, special characters, URLs, and HTML tags. After that, use CV and TF-TDF to convert numerical vectors from text review and then SVM to generate sentiment scores. Give this score to ANN for classification as positive or negative polarity [9].

Jaspreet Singh and Gurvinder Singh developed Aspect-based sentiment analysis. Take user reviews and preprocess for stop-word removal, POS tagging, NLP and ML approach to find QoS features, logistic regression to improve QoS of realistic data set. This model used a statistical method to analyze user reviews and assign positive and negative scores to them. For calculating the probability of QoS characteristics, use the logistic regression coefficient formula [10].

For sentiment analysis, Jaspreet Singh and Gurvinder Singh employed a Support vector machine (SVM) and Gradient boosting decision tree (GBDT) and discovered that SVM worked well when the data was simple but not so well when the database was sparse. For a complex database, GBDT functioned great, although it had an overfitting problem [10].

Gurvinder Singh and Jaspreet Singh work on the post-processing procedure. This model uses a pre-processed file as input before tagging POS with a text file loader (in Python 3.6) that uses UTF-8 encoding. This encoded token is assigned to a class for labeling and searching for QoS features. This model considers reviews of web services such as Amazon Cloud, Amazon Antivirus, and Amazon Software Services. [10].

Jeong Woong Sohn and Jin Ki Kim collaborated on a social network service (SNS) a web-based service that allows users to share their thoughts and experiences with their connections. Consumers can read and share other people's ideas and interests in a product on a range of venues, including product reviews, blogs, and social media sites, paradigm built the concept of word-of-mouth [11].

SNSes primarily deal with all product-related websites and analyze entire sites to measure reliability, trust, and web credibility. Content sharing, identity management, and context awareness are all functions provided by social networking sites. These services supply trustworthiness, e-commerce price, and cost, which are keys factors in making online purchasing decisions. The main characteristics of social commerce are economy, necessity, reliability, interaction, and sales promotion [11].

Jeong Woong Sohn and Jim Ki Kim's methodology extracts social commerce qualities that determine buy intention from shopping malls, homepage information, SNS functionalities, and social commerce features [11].

Valerio Basile et al. developed SENTIPOLC (SENTIment POLarity Classification Task) using Italian tweets as part of the EVALITA evaluation campaign. This natural language processing system determines whether tweets are subjective, polarised, or ironic. Based on the content of tweets, this model classified them as subjective or objective. Polarity is also identified and utilised to categorise sentiment as positive, negative, neutral, or mixed. Last but not least, irony detection assesses if a communication is ironic or not. These sardonic remarks have the wrong polarity. The system's potential application for irony patterns is demonstrated by annotations such as literal polarity in these texts [12].

Valerio Basile and colleagues compare data sets and results from the former gold standard SENTIPOLC in 2014 and 2016 to the new gold standard SENTIPLOC in 2018. The most popular algorithms covered by this model include SVM, Navie Bayes, K-Nearest Neighbors, and Decision Trees. In this paradigm, native preprocessing tools including tokenizers, POS tagging, and parsers include SentiWordNet in AFINN, HU-Liu Lexicon, and Whissel's Dictionary. This paradigm employs word-based encoding, which is subsequently supplemented with syntactic and semantic information, as well as microblogging components like emoticons and hashtags. The procedure included the use of ColingLab, INGEEOTEC, and CoMoDi. Deep learning algorithms like UniPI and SwissCheese [12] have also found success with this method.

Valerio Basile et al. used external resources for sentiment analysis after exhausting internal resources. In this model for SA, we employed Samskara, ItaliaNLP, CoLingLab, CoMoDi, Unitor, and IRADASE. This model utilises CNN for labelling with SwissCheese and Unitor [12].

Cornelius Puschmann and Alison Powell used stories from USA Today, The New York Times, The Washington Post, The Guardian, and The Daily Telegraph to look into opinion mining. This data set includes trade publications like Advertising Age, Bank Technology News, and Institutional Investor [13].

Linguistic Inquiry and Word Count (LIWC), developed by Cornelius Puschmann and Alison Powell, calculates the degree among different people who utilize various types of words in a text, emails, speeches, poetry, translation, and transcript.

Cornelius Puschmann and Alison Powell worked on identifying aspects for approximating human judgment using sentiment analysis. This model primarily comprehends sentiment

analysis and other analytic methodologies to comprehend social media. This strategy focuses on obtaining more public debate from social media and using computational methods to analyze and evaluate the information, as well as effectively forecast "user behavior" [13].
Mirsa Karim and Smija Das worked on sentiment analysis using machine learning and rule-based mechanisms. This model uses a dataset of movie reviews classified with positive and negative labels from Cornell University. This technique is a rule-based mechanism that employs mining algorithms to identify product features and opinions [14].
The Valence Aware Dictionary and Sentiment Reasoner (VADER) model used by Mirsa Karim and Smija Das calculate the compound score. More than and equal to 0.5 is considered good, less than and equal to -0.5 is considered negative, and -0.5 to 0.5 is considered neutral based on the compound score [14].
The sentiment analysis was handled by Smija Das and Mirsa Karim. Client ratings and reviews were used as input, and text mining was used to perform sentiment analysis. To compute the compound score, POS tagging, and synset score for vocabulary creation and training, this model used LDA analysis for NAVE BAYES, sentiment VADER, and sent word net techniques. The main purpose of this model is to provide a method for predicting the end of a sentence that is more precise and accurate.

METHODLOGY

Binary and multi-class classification approaches are two types of sentiment classification techniques. Using binary sentiment classification, the document-based model was separated into two labels, positive polarity and negative polarity. The texts are divided into numerous classifications, such as very positive, positive, neutral, negative, and severely negative polarity [9].

Figure 02 depicts the general architecture of sentiment analysis. The primary purpose of this research is to collect customer product reviews in the form of tweets. To ensure that the dataset is clean, certain data preparation is required, such as stop words removal, stemming, lemmatizing, chink, and chunking.
After preprocessing, sends this dataset to POS tagging with Navie Bayes Method or WordNet with Navie Bayes Method. Machine learning algorithms such as regression and SVC can now use the data set. Pickles are made to improve model speed by removing the need to create an environment for changing the machine learning strategy repeatedly. Following this concept, all results from various machine learning techniques are shown on a graph so that methodologies and analytical models may be compared.
It's a model of supervised machine learning. The data was gathered from Twitter and used to create a ternary classification system. This classification is divided into three categories: positive, negative, and neutral. This model employs ternary classification to classify given reviews using labeled data from a dataset. This model also compares two methods and provides the results in terms of accuracy.

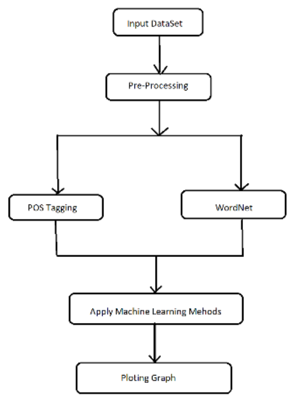

Fig. 2.Generic Architecture for Text Analytics by NLP

Because tweets are limited to 140 characters, this is the primary justification for using them as a data set. However, this data set requires additional preprocessing, removing the hashtag '#,' at the rate '@,' and retweeting. It imports the cleaned data collection's extracted data. A POS tag has been added to this data set manually. Before delivering information to the POS tagger, this model used nltk tokenizers such as sent tokenize and word tokenizer. Extract features from the dataset and divide it into two groups: training and test.

CONCLUSION

Customers analyzed reviews before purchasing any product. Customers read reviews regarding products, companies, services, and brands to gather information before purchasing. The review-based analyses open up the larger arena for research. The merchant can predict the customer's viewpoint based on reviews. By applying document-based sentiment analysis, customer evaluations can help to accurately forecast. Gather all document-based sentiment analysis concerns and challenges that may be classified as binary or multiclass. Propose work based on ternary classification and document-based sentiment analysis. That divides documents into three categories: good/negative/neutral.

REFERENCES

1. B. Seetharamulu, B. Naresh Kumar Reddy and K. Bramha Naidu, “Deep Learning for Sentiment Analysis Based on Customer Reviews”. IEEE 11th ICCCNT, 2020[ICFAI foundation for higher Education, Hyderabad, India].
2. Azwa Abdul Aziz, Andrew Starkey, Elissa Nadia Madi, “Prediction supervised Machine Learning Performances for sentiment analysis using contextual based approaches”, IEEE Access, 2019.
3. Karthikkayini T, N. K. Srinath. “Comparative polarity analysis on Amazon product reviews using existing machine learning algorithms”. IEEE International Conference, 2017.
4. Ronen Feldman. “Techniques and Application for Sentiment Analysis”. ACM, vol. 56, no. 4, 2013.
5. K. Saranya and Dr. S. Jayanthy, “Onto-based sentiment classification using Machine Learning Techniques”, ICIIECS, 2017.

6. Nadjim Frechet, Justin Savoie and Yannick Dufresne, "Analysis of Text-Analysis Syllabi: Buliding a Text-Analysis Syllabus Using Scaling", American Political Science Association, 2019.
7. K. Sudheer, Dr. B Valarmathi, SITE, "Real Time Sentiment Analysis of E-Commerce Websites using Machine Learning Algorithms". IJMET, Volume 9, Issue 2, February 2018, pp 180-193.
8. Lijuan huang, Zixin Dou, Yongjun Hu and Raoyi Huang, "Textual Analysis for Online Reviews: A Polymerization Topic Sentiment Model". IEEE Access, Volume 7, 25 July 2019, pp 91940-91945.
9. Abinash Tripathy, Abhishek Anand and Santanu Kumar Rath, "Document-level sentiment classification using hybrid machine learning approach", Springer-Verlag London 10 May 2017.
10. Jaspreet Singh and Gurvinder Singh," Sentiment Analysis of Social Media Reviews using QOS Parameterization". ICSCCC, 2018.
11. Jeong Woong Sohn, Jin Ki Kim, "Factors that influence purchase intentions in social commerce". ScienceDirect, 26 August, 2020.
12. Valerio Basile, Nicole Novielli, Danilo Croce, Francesco Barbieri, Malvina Nissim and Viviana Patti, "Sentiment Polarity Classification at EVALITA: Lessons Learned and Open Challenges". IEEE, 2018.
13. Cornelius Puschmann and Alison Powell, "Turning Words Into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media". SAGE, Social Media + Society July-September 2018.
14. Mirsa Karim and Smija Das, "Sentiment Analysis on Textual Reviews". IOP publishing, conf: Series Material Science and Engineering 396, 2018.
15. Jin Zhou and Jun-min Ye, "Sentiment analysis in education research: a review of journal publications". Interactive learning environments. 12 September, 2020.