Sentimental Analysis of Product Reviews with combined CNN-LSTM model using Deep Learning

[1]Pagadala Chinmayi
Department of Information Technology,
Vardhaman College of Engineering
Hyderabad, India
chinmayipagadala1@gmail.com

[2]Ganesh Bhaiyya Regulwar [0000-0001-9320-7040]
Department of Information Technology,
Vardhaman College of Engineering
Hyderabad, India
ganeshregulwar@gmail.com

[3]Navghare Anoop
Department of Information Technology,
Vardhaman College of Engineering
Hyderabad, India
navghareanoop281@gmail.com

[4]Karnati Geetha Reddy
Department of Information Technology,
Vardhaman College of Engineering
Hyderabad, India
karnatigeethareddy9@gmail.com

[5]Pathlavath Sunitha
Department of Information Technology,
Vardhaman College of Engineering
Hyderabad, India
sunithashankar280@gmail.com

2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU) | 979-8-3503-6178-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/IC-CGU58078.2024.10530690

Abstract— **Automated textual data sentiment analysis is the main focus of sentiment analysis, a crucial area within NLP. This abstract gives a succinct review of the state of sentiment analysis today, focusing on methodology, difficulties, and prospects. Sentiment analysis has progressed significantly in recent times because of the explosion of user-generated material on digital platforms. Deep learning algorithms in particular have made it possible to extract sentiment from text, enabling insights into public sentiment across a variety of fields. Despite this, context dependence, sarcasm detection, and the requirement for large, domain-specific datasets present difficulties for sentiment analysis. The importance of ethical issues, particularly bias and privacy concerns has increased. This paper analyzes the uses of sentiment analysis in marketing, politics, and healthcare while charting the development of sentiment analysis methodologies.**

Keywords— Sentiment Analysis, Deep Learning, Word Embedding, Accuracy, LSTM, CNN.

I. INTRODUCTION

Social media is meant to bring people together, exchange knowledge and viewpoints with others, but businesses may also use it to connect, learn about, and improve their goods and services. The number of people using social media increased. In 2019, there will be 2.77 billion social media users are active. globally, according to estimates. Feeling analysis is a technique that takes opinions from texts and extracts, converts, and interprets them in order to categorize them as positive, negative, or natural emotion [1]. Machine learning and lexicon-based techniques are the two primary subcategories of sentiment analysis methods. The lexicon-based approaches to determining the orientation of text content estimate the word and phrase semantic orientation as determined by a dictionary. [2].

Different sentiment analysis methods have been investigated by researchers, including: i) supervised techniques such as Naive Bayes (NB) and Support Vector Machine (SVM), Maximum Entropy (ME), Random Forest (RF), and so forth; and ii) unsupervised methods such as SentiWordNet (SWN), Semantic Orientation - Pointwise Mutual Information - Information Retrieval (SO-PMI-IR), etc[3].With features from Bag-of-Words, the NB model is the best, with 82.82% test accuracy. Furthermore above 80% are the sensitivity and specificity, which show that there is impartiality regarding either the favorable or unfavorable class. The first alternate is features of the Naïve Bayes model weighing 1-2 gm that achieves a strikingly similar outcome.

The top SVM classification achieved 81.92% test accuracy by utilizing the features of Bag-of-Words as well.[4] Within this research, it is evident that NB models perform better than any type of Bag-Of-Word features and SVM models are regarded as the primary characteristic of this issue. Online data types have a number of shortcomings that could increase the difficulty of analysis. The initial problem is that people are free to contribute their own material, which makes it impossible to guarantee the quality of their thoughts. For instance, rather than debating pertinent points of view, internet on forums, spammers post malicious spam. [6] Spam is completely worthless, while other spam contains irrelevant or "fake" perceptions. The second issue is that actual truth for such internet information is not readily accessible.

Instead, it uses word embedding as source containing context knowledge, and the neural network's intermediary layers pick up the features on their own in the training stage [8]. Using a variety of methods for data vectorization, including BOW (bag-of-Words, Glove), before using various ML methods for categorization, such as Logistic Regression (LR), NB, and taking the trials using a multi-class classification algorithm [13].

II. LITERATURE SURVEY

In the realm of sentiment analysis and opinion mining, various authors have explored diverse datasets and methodologies to derive meaningful insights. Vipin Deep Kaur, in [3], employed the GoodReads Dataset and Amazon Book Reviews Dataset, utilizing the So-PMI-IR, SVM, and Naïve Bayes methods. Through 5 and 10 folds, the median accuracy of Naïve Bayes ranged from 73.72% to 74.73%, while SVM exhibited accuracies of 73.46% and 74.70%, respectively. The unsupervised method yielded an accuracy of 73.23%.

In another study by Xing Fang and Justin Zhan [6], product reviews from Amazon were analyzed using Naïve Bayes, Random Forest, and SVM. Notably, the NB and SVM models demonstrated equivalent performance, outperforming the Random Forest model across various vector sets. However, at the sentence-level classification, both models faced challenges due to their comparatively poor performance on the impartial group.

Gagandeep Kaur and Amit Sharma, in [7], explored sentiment analysis using the Sentiment140, STS-gold, and SemEval-2014 datasets. Their approach involved Supervised Aspect-Based Sentiment Analysis (SABSA), SentiVec, N-gram+TFIDF+SVM, SEML, MTMVN, and HFV+LSTM. Significantly, N-gram+TF-IDF+SVM and SentiVec outperformed ABSA algorithms (MTMVN, SABSA, and SEML) by not extracting review-specific properties and neglecting to handle negation.

Ghazi A and Fatih Ö, as outlined in [8], delved into real-time Twitter data for sentiment analysis. Their methodology incorporated CNN and Bi-LSTM, with Bi-LSTM achieving remarkable results across various metrics and attaining a maximum accuracy rate of 90.3%. On evaluation with 20k data, the model accurately predicted 18060 tweets while mistakenly forecasting 1940 tweets.

Lastly, Ashok Kumar D and Anandan Chinnalagu, in [10], focused on sentiment analysis of COVID-19-related data using BLSTM and SAB-LSTM. SAB-LSTM outperformed conventional models like BLSTM and LSTM, attributing its effectiveness to the network optimizer layer, model architecture, and word embedding layer. Incorporating dropout and dense layers further enhanced the model's accuracy during training, resulting in improved prediction accuracy.

Both supervised and unsupervised methods are used to analyze the sentiments expressed in book reviews. In order to achieve this, I have used the widely used methods, including Naïve Bayes, Support Vector Machine, and SO-PMI-IR on two databases which are extracted from Amazon and GoodReads.[3] The Bag-Of-Words technique is used to extract the features from each post. Using this technique, a vector that represents each word's frequency in a post is created. The training dataset's posts must be searched for every unique term before creating the word list.[4] A word list's length is equivalent to the length of a feature vector. Surprisingly, all words with a frequency lower than 10 are dropped from list of words.

The outcomes reveal demonstrated unsupervised algorithms produced superior outcomes in the lengthy phrases are present in the dataset, although supervised. On the dataset including, algorithms provide more accuracy one-line summaries of the books.

Dictionary-based methodology within a lexicon-based Method has been applied to machine learning methods. Every product review has undergone sentiment analysis, after which it was categorized using machine learning techniques like NB and SVM.[5] More like a label applied to a judgment stating whether it is neutral, ideal, or adverse, a ground truth is that. Deep learning techniques have recently demonstrated efficient performance in applications involving sentiment analysis across numerous datasets for natural language processing.

To train the classifiers, each input of training data must be transformed into a feature vector, or a vector comprising those features. [6] Based on a sentence (review), one creates a feature vector for the assess-level categorization. Managing each vector's dimensionality is one of the challenges. Supervised ABSA (SABSA): This technique uses the SemEval-2014 dataset to do sentiment analysis by assessing a proposed aspect category prediction using supervised machine learning model that makes use of the co-occurrence strategy. Since the SABSA method and in this ARF technique are closely related, as per need selected it to compare with the recommended model. [7] SentiVec : This recently introduced method performs word embedding for sentiment analysis using a hybrid approach that combines supervised and unsupervised machine learning techniques. The ability to do away with the necessity for manual feature extraction is the main advantage of deep learning.

The work centered on applying sentiment analysis techniques to categorize the opinions expressed in product evaluations as positive or negative. In the research procedure, in this used Bi-LSTM, CNN, and CNN-Bi-LSTM. Word embedding was done using the word2vec technique. In this study,[8] tweets were automatically gathered from Tweepy on Twitter, and all strategies were tested on a 200k dataset split up into 50 percent positive tweets along with 50 percent negative tweets. The outcome demonstrates that the model SAB-LSTM by the authors performs superior to the conventional LSTM and BLSTM models. The connections between the targets and the opinion terms A bootstrapped dependency parser is used to identify syntactic relations. [9] Opinion word seeds are employed in the initial opinion lexicon, and the procedure employs semi-supervised techniques. The initial opinion lexicon is used to begin the bootstrapping process. Information is transferred back and forth between opinion words and targets via the double propagation method. The models were able to avoid overfitting issues and dynamically improve the model specifications for the specified dataset with the addition of the authors' additional layers. [10] In comparison to conventional LSTM models, the SAB-LSTM result's validation demonstrates greater sentiment prediction based on context.

In this paper use the Bag-Of-Words approach to extract the features for each article. Using this technique, a vector that represents the how frequently a word appears in a post is created. The training dataset's posts must be searched for every unique term before creating the word list. A word list's length is equivalent to the length of a feature vector. Surprisingly, all words with a frequency lower than 10 are dropped from this list of words.

Product review analysis focuses mostly on text analysis to identify the polarity of the reviews. The reviews are the most time-consuming activity, thus user-written and without a set format or set of guidelines to follow mapped [11]. Double propagation is used in feature extraction so that multiple product features can be extracted from a single review sentence. Accuracy and execution time can be impacted by factors such as data volume, composition, and training and testing methods for data selection. Product reviews are categorized using techniques that are regarded as effective and include good classification. Based on experiments, the findings show that, when Naive Bayes and SVM are employed as the two classification techniques, the accuracy of the SVM approach is higher than that of the NBC method, and the SVM method takes less time to execute than the NBC method. [12] It is challenging to assemble data from several platforms and polarize it. The researchers offer data filtration, a method that compiles both favorable and unfavorable reviews into a single file. Since the file contains an excessive number of extraneous numbers or symbols, they remove all unnecessary characters to improve a model's performance [13].

Customers and business owners look for a platform that allows for face-to-face communication. Raw data typically contains a large number of undesirable components that have a direct impact on how well machine learning or deep learning model’s function. [14] Because of this, have eliminated those components from the dataset, including stop words, punctuation, and unnecessary characters. Framing is the term for when online sources employ text as deliberate devices to present salient characteristics and points of view about a topic, employing specific keywords and cast images and sentences to communicate implicit meanings about a topic. How an issue is framed can have an impact on public opinion, attitudes, beliefs, and behaviors, especially when it comes from leaders with the ability to manipulate the public in addition to just exerting influence. [23] The goal of this study is to diagnose sepsis early by analysing physiological data. Patients' vital signs, test results, and demographic data are collected as inputs. An LSTM and an SVM were used to select the best hyper parameters for the training phase and the probability threshold for inference.

[15] Scores and tokens representing sentiment are generated using data from the original dataset. These will be mentioned to as features and utilized in the sentiment categorization process. Each training data entry must be transformed into a feature vector, or a vector with those characteristics, so that the classifiers can be trained. Based on an evaluation of a sentence), a feature vector is produced for the sentence-level categorization. Managing each vector's dimensionality is one of the challenges. Actually, there are two difficulties at hand: Owing to the curse of dimensionality, for the classifiers to suit the data, every vector ought to have an equal number of dimensions. rather than thousands or even hundreds of features or feature values. [16].

Numerous techniques for sentiment analysis have been researched, as well as the various degrees of sentiment analysis [20]. In this paper ultimate goal is to develop Sentiment Analysis [22], which will effectively classify different types of reviews. This article included concise talk about several fascinating machine learning methods include NB, SVM, Maximum Entropy techniques that can improve the analysis process. A careful examination of the text's semantics is important [17]. Feature level sentiment analysis, as it is also known, contains feature-based opinion mining and summarization [18]. It is crucial to ascertain precisely what individuals liked or disliked about this level. It is a level of sentiment analysis that is more precise. Aspect level examines the viewpoint directly rather than examining documents or words. [19] With three classes, the system showed a maximum F-measure score of 75.45%. The research community can access the gathered training collection of words and their embeddings. Neural network inputs are pre-trained vectors from Word2Vec. [21] Neither sentiment lexicons nor manually created features are used in this method.

III. PROPOSED WORK

Sentimental Analysis is a deep chapter to analyze. In this paper many datasets and texts to vectorize and know the sentiment or the emotion behind it. In this paper like social media posts, conversational chats, twitter tweets, product reviews etc. In these, chose to deal with Product reviews to enhance the customer experience through the sentimental analysis.

In this opt to use CNN and LSTM models to implement analysis. CNN – Convolutional neural networks are one kind of neural network for deep learning design that is widely used in computer vision. The area within artificial intelligence referred to as "computer vision" gives computers the ability to comprehend and analyze pictures and further visual information. Likewise, LSTM - Long Short-Term Memory, or LSTM, is a well-liked neural network that recurs (RNN) architecture in deep learning. It is excellent at identifying long-term dependencies, which makes sequence prediction tasks a perfect fit for it.

The real agenda to choose them is their efficiency of working with large datasets and producing the best results. Though the Deep Learning has a wide variety of algorithms, CNN gives the efficient results with its applications and hyper parameter tuning. Figure 1 is the combined architecture of CNN-LSTM model used:

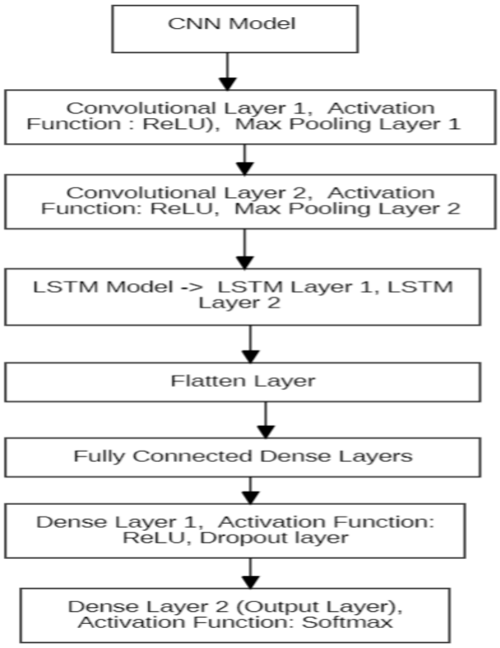

Figure 1. Architecture of CNN-LSTM model

IV. IMPLEMENTATION

Sentiment analysis involves the following steps: gathering data, pre-processing, word embedding, sentiment detection, and deep learning approach classification. Pre-processing turned the information into a regular text, enhancing it more machine-readable than it was in the original format. Word embedding is used in the following step to represent words in numerical form. Lastly, Test reviews are fed into the model, and these trained classifiers do the categorization work, finally categorizing the reviews as either good or negative. In the training data, each review is associated with a label for the class. Classifiers get this data after which it is used for training and learning. Architecture of process methodology is shown in Figure 2.

A. Preparation: Gathering a labeled dataset is the first phase of the sentiment analysis preparation procedure. The Sentiment Analysis dataset is available for use in this context. The two categories in this dataset—positive, negative and neutral—reflect the opinions expressed in the reviews.

B. Preprocessing: Special characters, punctuation, and HTML tags should all be eliminated from the text data. To do this, you can use Python's library. It is also necessary to convert the text data to lowercase in order to eliminate case sensitivity.

C. Tokenization: The text data must then be divided into words or subwords. The 'keras.preprocessing.text.Tokenizer' function in Keras can be used for this.

D. Padding: You must pad each review to the same length because they are all different lengths. The 'keras.preprocessing.sequence.pad_sequences' function in Keras can be used to accomplish this.

E. Creating the model: CNNs, or convolutional neural networks, or long short-term memory are two examples of deep learning models appropriate for this particular task. The length of the padded sequences should be the input layer. To predict sentiment, a sigmoid function ought to be used in the output layer.

F. Dataset used: The dataset used is from Kaggle.com named "Dataset-SA.csv". It is sentiment analysis dataset which has product reviews, summary, remarks etc columns of various products.

G. Flow Architecture of Methodology

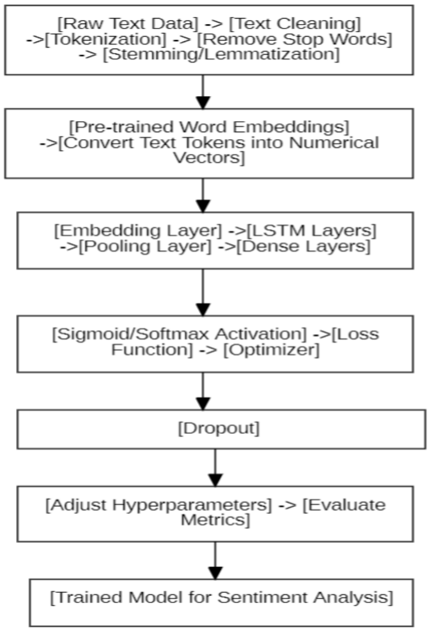

Figure 2. Architecture of Process Methodology

For better results, consider using a more sophisticated preprocessing method and applying techniques like hyperparameter tuning and regularization. Additionally, in this might want to use pre-trained word embeddings like Word2Vec or GloVe instead of training its own.

V. RESULTS

The general precision is 91.28% for the CNN-LSTM model. The dataset is classified into 2 groups as training and testing groups. The training data is trained with 10 epochs of 3862 values each. There isn't any significant difference within the different epoch's accuracy shown in Table 1. All are almost featured similar accuracy. The combined CNN-LSTM model worked better than the other CNN, LSTM and Bi-LSTM individual models. It also gave good performance in reviewing the product's analysis.

Table 1- Accuracy Comparison for different models

Model	*Accuracy*
CNN	79.36%
LSTM	90%
CNN-LSTM	91.28%

According to this dataset of about 30k rows, the model's performance was quite precise. But as need to move onto real-time data, the accuracy may differ due to inconsistent data shown in table 2. Confusion Matrix of the model which used is shown in Figure 3.

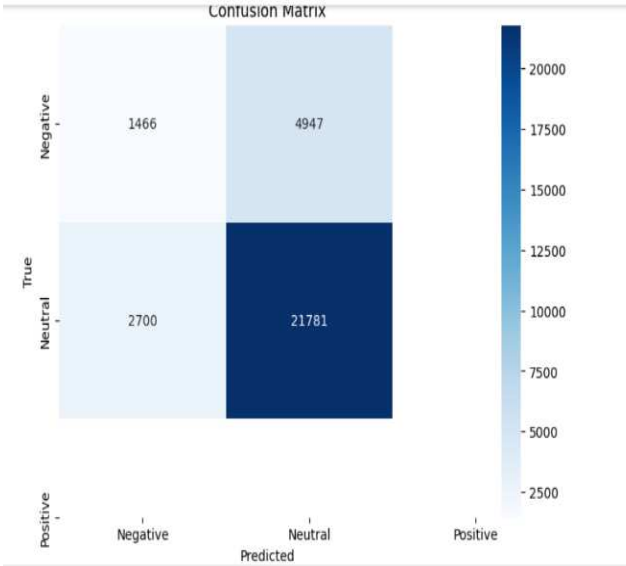

Figure 3. Confusion Matrix of the model

CLASSIFICATION REPORT:

Table 2 – Classification report of model's performance

	precision	recall	F1-score	support
0 (negative)	0.23	0.35	0.28	4166
1 (positive)	0.89		0.80.815	26728
Accuracy			0.75	30894
Macro Avg	0.56	0.58	0.56	30894
Weighted Avg	0.80	0.75	0.77	30894

VI. CONCLUSIONS AND FUTURE WORK

The model has given good accuracy compared to other models but however Bi-LSTM gave 90% of accuracy when used with the twitter datasets. And for product reviews dataset, CNN-LSTM works efficiently than any other models with 91.28% accuracy for Product reviews dataset.

When it comes to Future works, one can work on these,

i) Transfer Learning Approach which is applying a sentiment analysis model that has already been trained to the existing framework. By utilizing pre-existing knowledge from comparable tasks, this method enhances the performance of the model.

ii) Context-aware Analysis which is developing a model that takes into account the context of the words or sentences being examined. This will enable the model to predict sentiment more accurately and pertinently.

iii) Ensemble Learning Methods which is to increase the predictive performance of the model, use sophisticated ensemble learning strategies like bagging, boosting, and stacking.

REFERENCES

1. Drus, Z. and Khalid, H., 2019. Sentiment analysis in social media and its application: Systematic literature review. Procedia Computer Science, 161, pp.707-714
2. Onan, A., 2021. Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. Concurrency and Computation: Practice and Experience, 33(23), p.e5909.
3. Kaur, V.D., 2018, August. Sentimental analysis of book reviews using unsupervised semantic orientation and supervised machine learning approaches. In 2018 Second International Conference on Green Computing and Internet of Things (ICGCIoT) (pp. 519-524). IEEE.
4. Charoenkwan, P., 2018, July. ThaiFBDeep: a sentimental analysis using deep learning combined with bag-of-words features on Thai Facebook data. In 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI) (pp. 565-569). IEEE.
5. Jagdale, R.S., Shirsat, V.S. and Deshmukh, S.N., 2019. Sentiment analysis on product reviews using machine learning techniques. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2017 (pp. 639-647). Springer Singapore.
6. Fang, X. and Zhan, J., 2015. Sentiment analysis using product review data. Journal of Big Data, 2(1), pp.1-14.
7. Kaur, G. and Sharma, A., 2023. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. Journal of Big Data, 10(1), p.5.
8. Ramadhani, A.M. and Goo, H.S., 2017, August. Twitter sentiment analysis using deep learning methods. In 2017 7th International annual engineering seminar (InAES) (pp. 1-4). IEEE.
9. Jebaseeli, A.N. and Kirubakaran, E., 2012. A survey on sentiment analysis of (product) reviews. International Journal of Computer Applications, 47(11).
10. Kumar, D.A. and Chinnalagu, A., 2020, December. Sentiment and emotion in social media COVID-19 conversations: SAB-LSTM approach. In 2020 9th International Conference System Modeling and Advancement in Research Trends (SMART) (pp. 463-467). IEEE.
11. Dhabekar, S. and Patil, M.D., 2021. Implementation of Deep Learning Based Sentiment Classification and Product Aspect Analysis. In ITM Web of Conferences (Vol. 40, p. 03032). EDP Sciences.
12. Basani, Y., Sibuea, H.V., Sianipar, S.I.P. and Samosir, J.P., 2019, March. Application of sentiment analysis on product review e-commerce. In Journal of Physics: Conference Series (Vol. 1175, No. 1, p. 012103). IOP Publishing.
13. Haseeb, A., Taseen, R., Sani, M. and Khan, Q.G., 2023, April. Sentiment Analysis on Amazon Product Reviews using Text Analysis and Natural Language Processing Methods. In International Conference on Engineering, Natural and Social Sciences (Vol. 1, pp.

446-452).
14. Munna, M.H., Rifat, M.R.I. and Badrudduza, A.S.M., 2020, December. Sentiment analysis and product review classification in e-commerce platform. In 2020 23rd International Conference on Computer and Information Technology (ICCIT) (pp. 1-6). IEEE.
15. Jagdale RS, Shirsat VS, Deshmukh SN. Sentiment analysis on product reviews using machine learning techniques. InCognitive Informatics and Soft Computing: Proceeding of CISC 2017 2019 (pp. 639-647). Springer Singapore.
16. Fang, Xing, and Justin Zhan. "Sentiment analysis using product review data." Journal of Big Data 2.1 (2015): 1-14.
17. Devika, M.D., Sunitha, C[a]., Ganesh, A.: Sentiment analysis: a comparative study on different approaches. Proc. Comput. Sci. 87, 44–49 (2016)
18. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2004
19. Rosenthal, S., Farra, N., Nakov, P.: SemEval-2017 task 4: sentiment analysis in Twitter. In: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017) (2017)
20. M. Aarti and A. Patil, "Sentiment Analysis for Product Reviews," Int. J. Adv. Res. Comput. Sci., vol. 5, no. 5, pp. 202–204, 2014, doi: 10.21917/ijsc.2019.0266.
21. S.Smetanin and M. Komarov, "Sentiment Analysis of Product Reviews in Russian using Convolutional Neural Network," 2019 IEEE 21st conference on Business Informatics (CBI), IEEE, vol. 1, pp. 482-486, 2019
22. S.V. Balshetwar, G.B. Regulwar (June 2019) "Frame Tone and Sentiment Analysis", Published in International Journal of Computer Sciences and Engineering (IJCSE), E-ISSN: 2347-2693, Volume 7, Issue 6 pp.24-40, June 2019.
23. Rout, S.K., Sahu, B., Regulwar, G.B. and Kavididevi, V., 2023, January. Deep Learning in Early Prediction of Sepsis and Diagnosis. In 2023 (ICONAT) (pp. 1-5). IEEE.