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

Classifying Twitter Sentiment on Multi- Levels using A Hybrid Machine Learning Model

[1]Ananta Charan Ojha, [2]Pradeep Kumar Shah, [3]Sunil Gupta, [4]Sachin Sharma

Submitted: 21/08/2023 **Revised**: 08/10/2023 **Accepted**: 21/10/2023

Abstract: Social networking sites like Twitter have developed into rich sources for popular sentiment, providing insightful information about the attitudes, feelings, and reactions of the general public. In order to extract useful information from the enormous amount of content created by users on Twitter, sentiment analysis—the technique for autonomously recognising and classifying sentiment in textual data—is absolutely essential. In this article, we present a hybrid machine learning (ML)-based multilevel system for classifying Twitter sentiment. The Spiking neural network (SNN) is used in the suggested hybrid strategy (SNN+NB) suggested in this article in order to assist Naive Bayes (NB) classifiers make better decisions by giving NB an additional input. We employ a multilayer text mining application, involving data retrieval and query handling, to show the framework's functionality for language interpretation and sentiment analysis utilising hybrid machine learning. Performance metrics for our suggested SNN+NB method's accuracy, precision, recall, and F1-score are examined. The created framework can offer insightful information to people, organisations, and researchers looking to comprehend trends in public mood and make data-based choices using Twitter data.

Keywords: *Social network, Twitter, sentiment classification, tweets, hybrid machine learning (ML), SNN+NB*

1. Introduction

Social media sites like Twitter has developed into effective tools to conduct knowledge exchange, viewpoint assessment, and public sentiment monitoring in recent years. The enormous amount of content created by users on Twitter offers a special chance to comprehend the opinions of the general public on diverse subjects (Medhi et al. (2022)). Considering its potential use in market studies, managing brands, public reaction assessment, and other fields, emotion analysis—the process of instantly recognising the sentiment portrayed in text—has drawn a lot of emphasis (Hosen et al. (2021)).

However, there are a number of difficulties in categorising sentiment on Twitter. Firstly, tweets are known for being brief (just 280 characters), which frequently encourages the usage of slang phrases, designations, and spelling errors (Rust et al. (2021)).Following that, classification of sentiment on Twitter frequently involves looking at various levels of sentiment, such as determining the sentiment directed towards a specific person, place, or thing inside a tweet as well as the general feeling of the tweet. A more comprehensive understanding of the thoughts and attitudes stated on Twitter is made possible by these multi-level feelings (Wang et al. (2020)).

Pre-processing methods like encoding and developing are used to handle the unique characteristics of Twitter data, including emoticons, hashtags, and user mentions. Using feature engineering, pertinent linguistic features are extracted, such as n-grams, part-of-speech tags, and syntactic connections (Gopi et al. (2023)).The model combines deep learning models like neural networks based on convolution or Recurrent Neural Networks to handle the complexities and subtleties while conventional algorithms for machine learning like Support Vector Machines, Naive Bayes, or Random Forests are used to capture context-specific information and relationships. Techniques for ensemble learning are used to increase the overall accuracy of sentiment classification (Alroobaea, 2020). In this paper, we introduce a multilevel hybrid machine learning (ML) method for categorizing Twitter sentiment. The hybrid technique (SNN+NB) outlined in this article employs the Spiking neural network (SNN) to help Naive Bayes (NB) classifiers by providing NB with an additional input.

The remainder of this paper is arranged as follows: Part 2- related work, part 3- methods, Part 4- Results, and Part 5- conclusion with limitations and future scope.

2. Related Works

Mendon et al. (2021), created a framework to analyse users' sentiments on Twitter regarding natural catastrophes consuming data pre-processing methodologies and a blend of machine learning, statistical modelling, and lexicon-based methodology. In contrast to affinitive and hierarchical clustering, they select TFIDF and K-means for emotion categorization. To identify themes, employ Latent Dirichlet Allocation, a pipeline of Doc2Vec and K-means, followed by multi-level polarisation index categorization

[1]*Professor, Department of Computer Science and IT, Jain(Deemed-to-be University), Bangalore-27, India, Email Id: oc.ananta@jainuniversity.ac.in*
[2]*Assistant Professor, College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Email id: pradeep.rdndj@gmail.com*
[3]*Associate professor, School of Computer Science & System, JAIPUR NAITONAL UNIVERSITY, JAIPUR, India, Email Id: sunil95@rediffmail.com*
[4]*Assistant Professor, School of Engineering and Computer, Dev Bhoomi Uttarakhand University, Uttarakhand, India, Email Id: socse.sachin@dbuu.ac.in*

and its time sequence assessment. The categorization of feelings according to resemblance and polarisation indices and the discovery of themes among the subjects addressed on Twitter are the study's main conclusions.

Madani et al. (2020) suggested method is a brand-new hybrid strategy built on fuzzy logic, which has three crucial steps (Fuzzification, Regulation Inference/Aggregation, and defuzzification), and information regaining framework (IRS) concepts as well. It does this by determining the degree of semantic overlap between a tweet for classification and two viewpoint documentation implementing the WordNet dictionary. If they had a sizable dataset of tweets, then they choose to parallelize our work utilising the Hadoop framework with its distribution data storage (HDFS) and the MapReduce programming style to solve the computation time issue. The experiments conducted demonstrate that our technique superiors several different approaches from the literature, and by including fuzzy logic, they enhance the categorization outcomes.

Samuel et al. (2020) presented a methodological overview of two key machine learning (ML) methods for classification and contrast their performance in organising Coronavirus Tweets of various lengths. With the Naive Bayes technique, they obtain an excellent categorization efficiency of 91% for brief Tweets. This study describes connected with methodologies, consequences, potential, and constraints while offering perspectives on the evolution of the dread feeling towards the coronavirus.

Jain and Jain (2019) performed sentiment assessment and categorization of tweets using the hashtag #Renewable Energy. To categorise tweets into three groups, they used six different machine learning algorithms. They have performed classification both with and without the use of feature selection approaches. To choose fewer features from the dataset, they employed the CfsSubsetEvaluation and Information Gathering feature selection methods. The outcome of the approaches used in this study demonstrates that feature selection strategies improve the efficacy of emotion categorization. Support Vector Machine (Using PUK Kernel) and the CfsSubsetEvaluation choice of features method produce the best accuracy (92.96%).

Kumar and Sangwan (2019) analyses rumour recognition studies utilising several machine language (ML) approaches and describes the rumour detection procedure. The focus of this analysis involves one aspect of the category task in which they determine whether or not status for rumour provided web content. This work may be expanded to include a multi-level, fine-grained assessment where rumours may be identified as scams, erroneous or other types of erroneous for identifying rumours, a variety of cutting-edge and hybrid machine learning approaches, including fuzzy and neurofuzzy, can be applied.

Rodrigues et al. (2022) concentrates on the identification of streaming Twitter spam messages and analysing the sentiment of both archived as well as real-time tweets. Two distinct datasets, one for emotion analysis and the other for spam identification, were employed in the suggested methodology. They implemented various vectorization methods and contrasted the outcomes. Several machine learning and deep learning techniques are used to the stable record and live, in-the-moment tweets to carry out sentiment evaluation and spam identification.

Patel and Passi et al. (2020) used machine learning algorithms for performing sentiment analysis on a language dataset was presented in this research as an innovative approach. To provide greater precision data and to shrink the dataset, the input tweets were separated and processed before to analysis. The emotional content of tweets was further analysed using artificial intelligence algorithms like naive Bayes, SVM, random forests, and KNN, and the efficacy of the classification algorithms was evaluated for the precision and dependability of the determined sentiment analyses. The outcomes demonstrate the value and effectiveness of the methods implemented and discussed in the study.

Bazzaz et al. (2021) developed the SDE-RF algorithm—a mixture of the SMOTE and DE optimisation algorithm—was used on the RF classifier to address the issue of class inequalities in spam identification. In an unbalanced dataset, machine learning algorithms build a biased model with extremely poor precision since traditional methods of classification demonstrate a propensity to favour the vast majority classes.

Alfrjani et al. (2021) suggested a Hybrid Semantic Knowledgebase-Machine Learning technique is proposed where an innovative Domain Feature Recognition technique makes use of essential public Linked Open Data sources (DBpedia, IMDb) and a thorough understanding of the determined domain in order to enhance the accuracy and recollection of the domain feature identification task. By using a brand-new Domain Characteristic-Sentiment Affiliation technique and a developed language for sentiment for each individual domain feature, this approach also increases the precision with which the evaluations' emotions are calculated. At last, by using a new Opinion categorization algorithm based on fusing information derived from a generated semantic knowledgebase with a quantitative dataset, the hybrid technique improves the accuracy of opinion categorization on a multi-point scale.

3. Proposed Method

3.1. Dataset

The Protected Semiconductor Disorganised Sort-Class Dataset, first an example of the encrypted semiconductor Sort-Class dataset is shown in Table I. Following a specific engineer has finished the attribute selection procedure, the dataset was used. This dataset has triple input parameters, two of which are quantitative and one of which is qualitative. Testing flags that are allocated by a machine are categorical. This parameter has 2145 different standard deviation categories, with values varying from 1 to 214. Residual power and internal die fluctuations are denoted by the numerals Numerical_1 and Numerical_2, correspondingly. The range of the Numerical_1 value is between 186,720 and 316,830, whereas the range of the Numerical_2 value is between 185,760 and 301,800.

The outcome, which represents the class test findings, was presented as a categorical type such as "1" or "0." Particularly, "0" denotes the overwhelming number of or excellent units, while "1" denotes the minority or poor units. This dataset contains 286,000 samples that were gathered. 94% of the dominant class and 8% of the minority class make up this unbalanced dataset.

3.2. Pre-processing

Pre-processing describes the initial processes required to optimise and alter the raw Twitter dataset prior it can be utilised in analysis or machine learning applications. Certain important phases are often involved in the pre-processing process. In the beginning, it entails eliminating superfluous elements like URLs, hashtags, mentions, and special characters while keeping crucial content such as text. Second, it frequently entails standardising the text by making it all lowercase and getting rid of all punctuation. Tokenization, as the third component, involves dividing the text into separate words or tokens for additional analysis. In addition, stop words—commonly used words with little meaning—may be eliminated during pre-processing. In order to handle word changes and enhance consistency, it may also incorporate stemming or lemmatization, which reduces words to their fundamental or core form. Pre-processing, in general, prepares the Twitter dataset through translating it into a tidy and organised format, facilitating more efficient modelling and analysis.

3.3 Spiking neural network (SNN)

Spiking neural networks (SNNs) can be used to classify Twitter sentiment by simulating the actions of biological neurons. Word-integration methods are used to encrypt the words that make up each tweet. The Leaky Integrate-and-Fire (LIF) neuron, a type of spiking neuron, is used in the network, and it conveys sentiment activation by releasing action potentials. Increases in rates imply good mood, and lower rates indicate negative sentiment, according to the rate of stimulation or spiking frequency of output neurons. By applying gradient-based algorithms for learning like Spike Prop, training entails changing the weights between neurons. The SNN analyses word embedding’s from input tweets during inference and produces spiking activity in destination neurons, enabling sentiment analysis by examining spiking patterns. It is important to carefully consider computational difficulties and the accessibility of specialised frameworks like NEST or Brian when developing SNNs for sentiment categorization. SNNs provide an alternate strategy, however due to their superior performance and widespread application in natural language processing applications, traditional machine learning algorithms like CNNs and RNNs continue to be preferred options for sentiment analysis in Twitter data.

3.4 Naive Bayes classifier

A probabilistic-based classifier known as a Naive Bayes classifier was first presented in 1973. The classifier has performed well in a variety of applications. Numerous researchers has provided thorough explanations of some essential Naive Bayes methods characteristics. A Naive Bayes network's structure consists of variables, X_j i 3,2,1,...,n and output, G. The Naive Bayes model presupposes that each of these networks input is independent. In a Naive Bayes network, the relationship between the inputs and the outputs was fixed and unchangeable.

A probability estimation strategy serves as a learning mechanism in the Naive Bayes classifier to assess the likelihood of each class during the prediction phase. To generalize the categorization framework, a continuous-valued input parameter or a distribution of Gaussian values is typically selected.

$$l(y|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(y-\mu^2)}{2\sigma^2}\right) \quad (1)$$

Where l is the source of X_j value. The variance of the data, X_j, is 2, and the mean of the data, X_j, is μ and σ^2. Following that, the likelihood of each input X and its probability, B(X), supplied by a class of G g, can be determined using the following equation:

$$B(X) = b(G = g) \prod_{j=1}^{m} b(X_j = x_j | G = g) \quad (2)$$

The probability of each input, B $(X)_j$, is determined earlier than the prediction procedure using a maximum parameter.

$$\boldsymbol{X^{predict}} = \mathbf{arg}\, \boldsymbol{max_g}\, \boldsymbol{B(X)} \quad (3)$$

3.5 Hybrid SNN-Naive Bayes Classifier

Combining the benefits of Spiking Neural Networks (SNNs) and the Naive Bayes algorithm, a hybrid SNN-Naive Bayes classifier is a model used for classification. For the purpose of trying to increase classification performance, the hybrid technique tries to combine the time-dependent perception capabilities of SNNs with the statistical character of Naive Bayes. The first step in the procedure is to optimise the information being supplied, extract pertinent features, and then encode them as spikes sequences in the SNN component. By spreading spikes over its network, the SNN is able to capture the variations in time and trends in the data. The Naive Bayes method is concurrently developed by employing the initial vectors of features to calculate class-conditional probability. Additional input data is represented as spiking trains and delivered to the SNN during evaluation. Spiking-based imitations of the supplied samples are produced by the SNN after processing the spike trains. The Naive Bayes component is then given these spike-based descriptions. Implementing Bayes' rule, the Naive Bayes algorithm determines the following possibilities of the class labels given the spike-based representations.

The last categorization choice is then based on the greatest likelihood by the Naive Bayes component. The hybrid classifier can accurately classify data by combining Naive Bayes' probabilistic thinking with SNNs' ability to identify patterns of time dependence in data collected through spikes. When working with time-series data, the Hybrid SNN-Naive Bayes Classifier is especially helpful because accurate categorization depends on changing temporal patterns. It offers a thorough and potent method to categorization through integrating the advantages of both models.

4. Result and Discussion

In this section, we evaluate the performance of suggested method for sentiment classification in twitter. The existing methods are CA-SVM, CNN, LSTM analysed in this paper. The analysis includes precision (%), accuracy (%), recall (%), and F1-score (%).

The percentage of actual accurate forecasts among all correct predictions is the classifier's accuracy. In sentiment evaluation, precision is a statistic that measures how well the machine learning model detected the appropriate sentiment for a specific piece of text. Or, the instances in which the classifier successfully classified a section of text as positive (i.e., the sum of genuine positive predictions and incorrect positive predictions) can be expressed by dividing the entire amount of correct predictions by the actual number of accurate predictions.

$$Precision = \frac{True\ positive\ predictions}{(True\ positive\ predictions\ + False\ positive\ predictions)} \quad (4)$$

Precision for the recommended and standard approaches is depicted in Figure 1. The suggested techniques SNN+NB receive 98% of the precision, whereas CA-SVM, CNN, and LSTM only get 90 %, the 70 %, and 75 %, respectively. Compared to conventional approaches, SNN+NB technique has a higher precision Percentage.

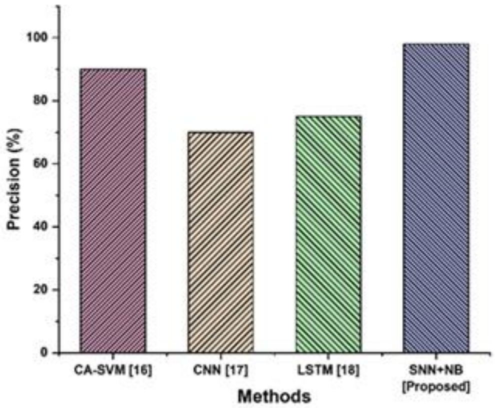

Fig 1: Precision

Accuracy is a statistic that is also used in sentiment analysis; it measures how well a classifier was able to identify the overall sentiment of a given text. The percentage of instances in the dataset that were appropriately categorised is known as accuracy shown in Figure 2. In this context, accuracy is calculated by dividing the total number of instances in the dataset by the number of instances that were effectively classified (i.e., the amount of instances the machine learning algorithm correctly identified a linguistic element as either positive or negative).

$$Accuracy = \frac{(True\ positive\ predictions + True\ negative\ predictions)}{Total\ number\ of\ instance} \quad (5)$$

The second figure shows the accuracy for the suggested and accepted methods. While CA-SVM, CNN, and LSTM only achieve 90%, 82%, and 70% accuracy, respectively, the suggested technique SNN+NB achieves 97% accuracy. The accuracy % of SNN+NB procedures is higher than that of conventional methods.

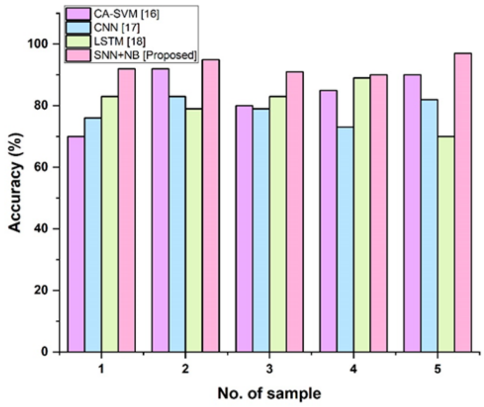

Fig 2: Accuracy

In sentiment analysis, the recall statistic is used to measure how well an algorithm is able to recognise every single instance of a particular sentiment within a dataset. Out of all the real-life favourable predictions, recall calculates the proportion of events that belong to a particular sentiment class. The number of times the machine learning algorithm correctly classified an element of text as either favourable or unfavourable (i.e., a combination of genuine optimistic predictions and erroneous adverse predictions) is used to calculate the recall, which is calculated by splitting the total number of instances that belong to a given sentiment class by the number of accurate positive forecasts that were made.

$$Recall = \frac{True\ positive\ predictions}{(True\ positive\ predictions + False\ negative\ predictions)} \quad (6)$$

Figure 3 illustrates a recall of the advised and accepted methods. The strategies recommended While CA-SVM, CNN, and LSTM only receive 81%, 90%, and 70% of the recall, respectively, RO-SNN+NB receives 97% of it. The Recall percentage of RO-TGANN procedures is higher than that of conventional methods.

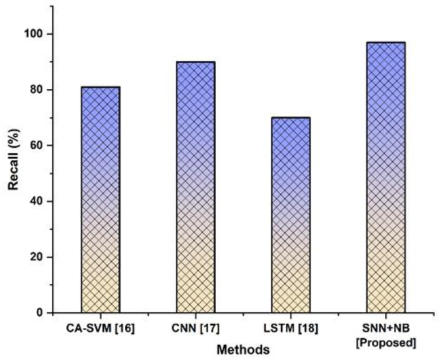

Fig 3: Recall

The F1-score is a parameter applied to sentiment analysis to determine how well a machine learning model performs overall by combining its precision and recall. The F1-score, which is a measurement that takes into account both erroneous positives and erroneous negatives, is the inverse mean of recall and accuracy.

$$F1 - score\ =\ 2\ *\ \frac{(precision * recall)}{(precision + recall)} \quad (7)$$

Figure 4 shows how the suggested and accepted methods are remembered. The strategies recommended In contrast to CA-SVM, CNN, and LSTM, which only obtain 81%, 82%, and 70% of the F1score, respectively, SNN+NB receives 96% of the score. The F1score percentage of SNN+NB procedures is greater than that of conventional methods.

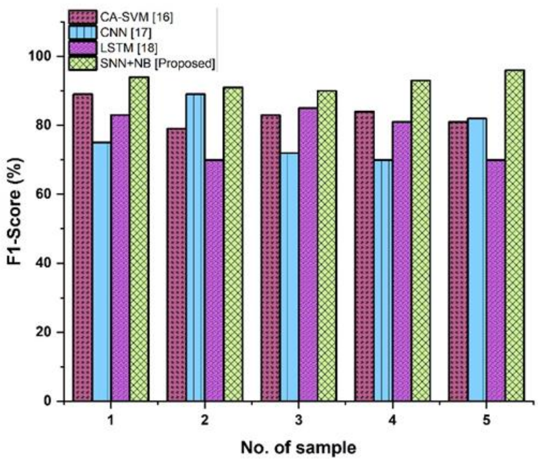

Fig 4: F1-Score

5.Conclusion

This study offers a hybrid machine learning-based multilevel system for sentiment analysis of Twitter data using a mixture of Spiking Neural Network (SNN) and Naive Bayes (NB) classifiers. The suggested SNN+NB technique improves NB's ability to make decisions by adding more information from SNN. The framework comprises a variety of layers of text mining application to show its capabilities in sentiment analysis while managing queries, interpreting language, and retrieving data. For the SNN+NB approach, performance indicators like accuracy 97%, precision 98%, recall 97%, and F1-score 96% are assessed. Overall, this paradigm provides insightful information to those who want to comprehend public mood patterns and make data-driven decisions using Twitter data, including individuals, organisations, and researchers. In future, we can implement this method in various social media such as YouTube, Instagram, Facebook and Amazon and so on.

References

[1] Hosen, M., Ogbeibu, S., Giridharan, B., Cham, T.H., Lim, W.M. and Paul, J., 2021. Individual motivation and social media influence on student knowledge sharing and learning performance: Evidence from an emerging economy. Computers & Education, 172, 104262.

[2] Wang, Z., Ho, S.B. and Cambria, E., 2020. Multi-level fine-scaled sentiment sensing with ambivalence handling. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 28(04), 683-697.

[3] Alroobaea, R., 2020. An empirical combination of machine learning models to enhance author profiling performance. International Journal, 9(2).

[4] Medhi, K., Hoque, N., Dutta, S.K. and Hussain, M.I., 2022. An efficient EEG signal classification technique for Brain–Computer Interface using hybrid Deep Learning. Biomedical Signal Processing and Control, 78, 104005.

[5] Rust, R.T., Rand, W., Huang, M.H., Stephen, A.T., Brooks, G. and Chabuk, T., 2021. Real-time brand reputation tracking using social media. Journal of Marketing, 85(4), 21-43.

[6] Gopi, A.P., Jyothi, R.N.S., Narayana, V.L. and Sandeep, K.S., 2023. Classification of tweets data based on polarity using improved RBF kernel of SVM. International Journal of Information Technology, 15(2), 965-980.

[7] Mendon, S., Dutta, P., Behl, A. and Lessmann, S., 2021. A hybrid approach of machine learning and lexicons to sentiment analysis: Enhanced insights from twitter data of natural disasters. Information Systems Frontiers, 23,1145-1168.

[8] Madani, Y., Erritali, M., Bengourram, J. and Sailhan, F., 2020. A multilingual fuzzy approach for classifying Twitter data using fuzzy logic and semantic similarity. Neural Computing and Applications, 32, 8655-8673.

[9] Samuel, J., Ali, G.M.N., Rahman, M.M., Esawi, E. and Samuel, Y., 2020. Covid-19 public sentiment insights and machine learning for tweets classification. Information, 11(6), 314.

[10] Jain, A. and Jain, V., 2019. Sentiment classification of twitter data belonging to renewable energy using machine learning. Journal of information and optimization sciences, 40(2), 521-533.

[11] Kumar, A. and Sangwan, S.R., 2019. Rumor detection using machine learning techniques on social media. In International Conference on Innovative Computing and Communications: Proceedings of ICICC 2018, Volume 2 (pp. 213-221). Springer Singapore.

[12] Rodrigues, A.P., Fernandes, R., Shetty, A., Lakshmanna, K. and Shafi, R.M., 2022. Real-time twitter spam detection and sentiment analysis using machine learning and deep learning techniques. Computational Intelligence and Neuroscience, 2022.

[13] Patel, R. and Passi, K., 2020. Sentiment analysis on twitter data of world cup soccer tournament using machine learning. IoT, 1(2), 14.

[14] Bazzaz Abkenar, S., Mahdipour, E., Jameii, S.M. and Haghi Kashani, M., 2021. A hybrid classification method for Twitter spam detection based on differential evolution and random forest. Concurrency and Computation: Practice and Experience, 33(21), e6381.

[15] Alfrjani, R., Osman, T. and Cosma, G., 2019. A hybrid semantic knowledgebase-machine learning approach for opinion mining. Data & Knowledge Engineering, 121, .88-108.

[16] Cyril, C.P.D., Beulah, J.R., Subramani, N., Mohan, P., Harshavardhan, A. and Sivabalaselvamani, D., 2021. An automated learning model for sentiment analysis and data classification of Twitter data using balanced CA-SVM. Concurrent Engineering, 29(4), 386-395.

[17] Alharbi, A.S.M. and de Doncker, E., 2019. Twitter sentiment analysis with a deep neural network: An enhanced approach using user behavioral information. Cognitive Systems Research, 54,50-61.

[18] Bisht, A., Singh, A., Bhadauria, H.S., Virmani, J. and Kriti, 2020. Detection of hate speech and offensive language in twitter data using lstm model. Recent trends in image and signal processing in computer vision, 243-264.

[19] Gummadi, A. ., & Rao, K. R. . (2023). EECLA: A Novel Clustering Model for Improvement of Localization and Energy Efficient Routing Protocols in Vehicle Tracking Using Wireless Sensor Networks. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2s), 188–197. https://doi.org/10.17762/ijritcc.v11i2s.6044

[20] Alexei Ivanov, Machine Learning for Traffic Prediction and Optimization in Smart Cities , Machine Learning Applications Conference Proceedings, Vol 3 2023.

[21] Gupta, R., Mane, M., Bhardwaj, S., Nandekar, U., Afaq, A., Dhabliya, D., Pandey, B. K. Use of artificial intelligence for image processing to aid digital forensics: Legislative challenges (2023) Handbook of Research on Thrust Technologies? Effect on Image Processing, pp. 433-447.