FAKE REVIEW DETECTION

LITERATURE REVIEW

G. M. Shahariar et al. focus on detecting deceptive reviews [1]. Researchers have made use of both labelled and unlabelled data and developed deep learning algorithms for spam review detection, including the Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), and a variation of the Recurrent Neural Network (RNN) known as Long Short-Term Memory (LSTM). Additionally, they used some classic machine learning classifiers to detect spam reviews, such as Naive Bayes (NB), K Nearest Neighbour (KNN), and Support Vector Machine (SVM), and compared the performance of both traditional and deep learning classifiers.

In [2], Khan H et al. propose using a supervised machine learning technology called support vector machine (SVM) to distinguish between fake (spam) and authentic (ham) text. The supervised learning technique for classifying spam and genuine reviews begins with entering the input review, pre-processing it, and then using the SVM classifier to classify it as fraudulent (SPAM) or genuine (HAM) on fake benchmark reviews. The proposed supervised machine learning technique for detecting false reviews is compared to previous work and other supervised machine learning classifiers.

J. T. Rodrigues et al. [3] suggested a method for assessing review feedbacks using deep learning neural networks such as the Gated Recurrent Unit (GRU), Bidirectional LSTM (Bi-LSTM), and Long Short Term Memory (LSTM), and evaluating the outcomes using activation functions such as ReLu, TanH, and Sigmoid. The dataset used to detect false product reviews and train suggested models was obtained from GitHub and consists of text reviews from customers on mobile phone products. This dataset consists of negative and positive reviews that have already been labeled and separated into two files. The text reviews are tokenized by breaking them down into smaller lines or words. The json file is loaded in the code to load the data, and only two data attributes, reviews and summary attributes, are used in the training of the data. Three models, LSTM, Bi-LSTM, and GRU, are utilized to define the model, each with a different activation function to see which one offers us the highest accuracy.

The main goal of the paper by S. M. Anas and S. Kumari [4] is to combat spammers by building a sophisticated model on millions of reviews. In this research, they used the "Amazon Yelp dataset" to train the models. Its short dataset is used for training on a small scale, but it may be scaled up to achieve great accuracy and flexibility. The fake review dataset is trained using two Machine Learning (ML) models that estimate the accuracy of how authentic the reviews in a dataset are. When relying on product reviews for items discovered online on various websites and applications, the rate of fraudulent reviews in the E-commerce business and even other platforms is increasing. This model may be used by websites and applications with tens of thousands of users to forecast the legitimacy of reviews, allowing website owners to take appropriate action.

I. Amin and M. Kumar Dubey [5] focus on malicious identification of views or feedbacks in the comments. The spam detection framework has improved efficiency over traditional machine learning and works well with unique models, providing better answers to challenging real-life scenarios, thanks to a surge in the effective uses of Soft computing. Most current research on review spam detection uses the standard bag-of-words model to recognize text analysis characteristics and to use traditional machine learning models such as Support Vector

Machines and other such classifiers. They have three proposed methodologies for this namely Feature Selection based on Linguistic Approach, Feature Selection based on Behavioral Approach of Reviewers related with its Products and lastly, Soft Computing Techniques.

Barbara Probierz et al. notice that a severe problem is a disinformation in the form of fake news. The researchers [6] have done some initial news analysis by its title. The purpose of the researchers was to create a novel model for initial news analysis and rapid detection of false news based solely on the headline rather than analyzing the complete article content. Furthermore, the ability to balance precision and recall as categorization quality measures was developed, allowing for better news selection. After analyzing the text with NLP approaches, the researchers have used the Adaptive goal function of ant colony optimization algorithms to find fake news. This research's hypothesis was that employing the goal-oriented ACDT method and a confined term matrix allowed for better classification than traditional algorithms. Experiments have shown that it is possible to conduct a preliminary analysis of fake and authentic news based on the collected data. Experiments showed that using a constrained word matrix, it is possible to undertake a preliminary examination of fake and genuine news. Adopting goal-oriented ACDT by the researchers has allowed them to significantly increase recall of real news and precision of fake news. Using decision tables will help the researchers gain better results, and increased model accuracy. Researchers think it is also worth looking into the impact of the number of words in the choice table on the outcomes and classification coverage. Following a comprehensive examination, it may be worthwhile to construct a two-stage verification system in the future – the first based on a restricted number of words from the news headline, and the second based on the complete content, but only for news that could not be categorized earlier. The method given here can identify false news by looking at the titles of stories that surface on the internet.

Rozita Talaei Pashiri et al. propose a feature selection-based method based on the sine–cosine algorithm presented to reduce spam detection inaccuracy (SCA). Feature vectors are updated by the proposed technique [7]. The precision, accuracy, and sensitivity of the proposed technique for the Spambase dataset in MATLAB were 98.64 per cent, 97.92 per cent, and 98.36 per cent, respectively. In other words, when it came to spam identification, the proposed method outperformed the multilayer perceptron (MLP) neural network, Bayesian network, decision tree, and random forest classifiers. The feature selection error in the MLP neural network was reduced by approximately 2.18 per cent employing the SCA, according to the test findings.

The spam detection error is affected by several elements: the ANN inputs and the selection of crucial attributes that best represent spam and emails. Because each feature has varying relevance, using all of them for training the ANN raises the inaccuracy. On the other hand, including all features increases both the problem and data size, increasing execution time. The results of the testing revealed that the suggested technique outperformed other learning methods in terms of accuracy, precision, and sensitivity in spam detection, including the MLP neural network, Bayesian network, decision tree, and random forest. In this aspect, the MLP neural network came in second. It is suggested that the proposed algorithm's detection error be reduced by upgrading it.

K. Archchitha and E.Y.A. Charles has proposed that CNN is offered to distinguish genuine opinion reviews from opinion spam. The proposed model was trained to utilize opinion text represented as word vectors by the pre-trained GloVe word embedding model. [8] This method varies from prior methods in two respects. Most earlier efforts relied on classifiers such as Support-Vector Machines and Nave Bayes to describe review text characteristics using the classic bag-of-words paradigm. The suggested CNN model outperformed existing techniques when evaluated on the Deceptive opinion spam corpus. The trials demonstrated that additional factors other than textual semantics must be investigated to identify deceptive viewpoints in reviews successfully. The suggested model confirmed its capabilities by enhancing accuracy even more. The suggested model's performance may be enhanced by employing a more significant data set, adding new characteristics such as behavioural information, and fine-tuning the CNN model's hyper-parameters. The classification method is similar to that used in many text-based machine learning applications. As a result, this model may also be used for sentiment analysis, auto-tagging of client questions, and text segmentation into predetermined subjects.

Ting-You Lin et al. have created a framework for detecting fraudulent reviews and conducted a thorough analysis on the efficacy of various variables and their combinations using three classifiers. [9] A diverse selection of popular benchmark datasets from Amazon review data and Yelp review data was used in the simulated trials. It has been discovered that readability and subject characteristics are more successful than sentiment analysis (sentiment features) in detecting bogus reviews. The essential feature group emerged as readability features.

When paired with FOG or FK's readability characteristics, topic features become even more crucial. The researchers are investigating other significant elements for detecting phoney reviews and comparing them to other research studies. Researchers have suggested and investigated making use of deep learning models.

Petr Hajek et al. suggested two deep NN models for identifying fake consumer reviews in [10], utilizing an integrated framework of n-gram, Skip-Gram, and emotion models. The experimental findings on four real-world fake review datasets in this paper indicate the efficacy of the suggested models. Notably, the suggested models beat current baseline techniques and state-of-the-art fake review identification systems in terms of accuracy, AUC, and F-score. The experimental findings of the researchers also revealed that the suggested integrated models were the most successful for more enormous datasets with coupled polarity, hinting that they might be used in real-world circumstances. The suggested detection methods were also successful in terms of time complexity and detection time, which the researchers revealed using the ARR multicriteria measure.

It was suggested that future research should integrate the presented models with graph-based techniques based on review information. In the future, researchers want to combine the benefits of the DFFNN and CNN models to create a hybrid deep NN structure akin to the Network in Network. A hybrid model like this might help the CNN model generalize even more. Another shortcoming of the suggested approach is that sentence weights were omitted due to their domain-specific character, unlike the CNN model.

REFERENCES

[1] G. M. Shahariar; Swapnil Biswas; Faiza Omar; Faisal Muhammad Shah; Samiha Binte Hassan, "Spam Review Detection Using Deep Learning," in *IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2019.

[2] Khan H., Asghar M.U., Asghar M.Z., Srivastava G., Maddikunta P.K.R., Gadekallu T.R, "Fake Review Classification Using Supervised Machine Learning," in *International Conference on Pattern Recognition*, 2021.

[3] P. Shetgaonkar, J. T. Rodrigues, S. Aswale, V. L. K. Gonsalves, J. C. Rodrigues and A. Naik, "Fake Review Detection Using Sentiment Analysis and Deep Learning," in *International Conference on Technological Advancements and Innovations (ICTAI)*, 2021.

[4] S. M. Anas and S. Kumari, "Opinion Mining based Fake Product review Monitoring and Removal System," in *6th International Conference on Inventive Computation Technologies (ICICT)*, 2021.

[5] I. Amin and M. Kumar Dubey, "An overview of soft computing techniques on Review Spam Detection," in *2nd International Conference on Intelligent Engineering and Management (ICIEM)*, 2021.

[6] Probierz B., Kozak J., Stefański P., Juszczuk P., "Adaptive Goal Function of Ant Colony Optimization in Fake News Detection," in *International Conference on Computational Collective Intelligence*, 2021.

[7] Rozita Talaei Pashiri, Yaser Rostami & Mohsen Mahrami , "Spam detection through feature selection using artificial neural network and sine–cosine algorithm," in *Math Sci 14*, 2020.

[8] K. Archchitha, E.Y.A. Charles, "Opinion Spam Detection in Online Reviews Using Neural Networks," in *IEEE*, 2019.

[9] Ting-You Lin, Basabi Chakraborty, Chun-Cheng Peng, "CNN is offered as a method for distinguishing genuine opinion reviews from opinion spam.," in *International Conference on Data Analytics for Business and Industry (ICDABI)*, 2021.

[10] Petr Hajek, Aliaksandr Barushka & Michal Munk , "Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining," in *Neural Comput & Applic 32, 17259–17274 (2020)*, 2020.