International Journal of Informatics and Computation (IJICOM)
Vol. 5, No.2, December 2023
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

A Comparative Study of Detecting Twitter Spam using Deep Learning

M Amman Said[1], Yaser Ahmad[2]

Abstract

This study addresses the escalating challenge of Twitter spam detection by leveraging the power of Convolutional Neural Networks (CNNs). With the proliferation of spam content on social media platforms, traditional machine learning algorithms have exhibited limitations in discerning intricate patterns within sequential data. The research problem centers on the need for a more robust and effective approach to distinguish spam tweets from legitimate content. The primary objective is to evaluate the performance of CNNs in comparison to baseline algorithms, including SVM, Decision Tree, KNN, Gaussian Naive Bayes, and Gradient Boosting. The research approach involves thorough data preprocessing, followed by model training and assessment using metrics like Confusion Matrix and Classification Report. The outcomes indicate that the CNN model outperforms the baseline algorithms, exhibiting superior levels of accuracy, precision, recall, and F1-score. These results highlight the promise of CNNs in reshaping the landscape of Twitter spam detection, presenting a more precise and effective approach to tackle the spread of spam content across social media platforms. This research contributes valuable insights for the development of advanced machine learning techniques in the domain of online security and spam detection.

Keywords

Twitter, Spam Detection, Convolutional Neural Networks (CNN), Deep Learning

1. Introduction

The significance of detecting spam on widely-used social media platforms such as Twitter cannot be overstated, given the immense popularity of these platforms and the growing occurrence of spam. Social media has evolved into a vital medium for communication and the dissemination of information. Individuals utilize social networking sites to connect with new individuals and engage in conversations with current contacts, and among these platforms, social media is experiencing the most rapid growth. Twitter, in particular, has emerged as a major source of spam, causing problems for users. With millions of active users and a large number of tweets posted daily, spammers are attracted to misuse the platform for their malicious activities**.** The surge in spam accounts on Twitter has led researchers to seek strategies to mitigate this problem. Effective spam detection methods are needed to filter out spam messages and create a spam-free environment for legitimate user [1].

The main purpose of the study in the context of spam detection on Twitter is to develop effective machine learning-based approaches for distinguishing spam messages from legitimate ones. The research suggests employing a range of attributes and classification methods, encompassing lexical, syntactic, and semantic attributes, along with decision trees, naive Bayes, and support vector machines have been utilized for this purpose. Identifying spam on social media platforms. [2]. The study also aims to compare the performance of different machine learning classifiers, including Decision Tree, Support

[1] Corresponding Author (Amman Said: amman.said@saba.edu), Saba University, Yemen
[2] Yaser Ahmad, Saba University, Yemen

Vector Machine, Naïve Bayesian, and Logistic Regression, regarding accuracy, robustness, and scalability, the aim is to pinpoint the most suitable classifier for detecting spam on Twitter [3]. Additionally, the study explores the use of ensemble machine learning ensemble models and majority voting models have been found to enhance the accuracy of predicting Twitter spam account detection in comparison to individual models. The study also evaluates the performance pertaining to two ensemble learning algorithms based on trees, namely Random Forest and Extra Trees, for Twitter spam classification [9]. In the introductory section, the problem is defined and expounded upon by scrutinizing the deficiencies and lacunae in the current body of literature pertaining to the subject matter. [4].

Numerous research endeavors have centered around identifying spam on the Twitter platform. In recent investigations, machine learning (ML) models have been employed to spot spam accounts on Twitter. It has been observed that utilizing ensemble ML models and employing a majority voting approach can enhance the accuracy of predicting Twitter spam account detection in comparison to individual ML models. Among the models assessed for identifying Twitter spam accounts with an imbalanced dataset, the Random Forest model emerged as the most effective choice. [20]. Furthermore, a Chimp Sailfish Optimization-based Deep Neuro Fuzzy Network (ChSO-based DNFN) has been introduced as a proficient approach for screening out spam content on various social media platforms, including Twitter [5].

A study proposed a conceptual matrix or hypothesis that combines various social problems in school education and relates them to the implementation of managerialism ideology. Convolutional Neural Network (CNN) was chosen as the lead algorithm in the study because its application is prevalent in the field of image processing and has achieved good classification results. CNN models were used to classify brain tumors based on various parameters. In the study on pneumonia diagnosis, a CNN model was built from the ground up and produced incredibly accurate results [6].

CNN has several advantages in processing sequential data such as text, especially in social media environments. Firstly, CNN can extract local information between consecutive words in a sentence, which helps capture important features and patterns in the text. Secondly, CNN can handle the spatial correlation of the data in each frame of the sequence, allowing it to capture the dependencies between frames [16]. Lastly, CNN can extract sentiment features with different granularity, enabling it to analyze the sentiment of social network texts effectively. These advantages make CNN a powerful tool for processing sequential data in social media environments, improving the precision of sentiment classification and enabling better prediction of emotional tendencies in short texts [7].

The study explored that ensemble models using cross-learning outperformed local time series Forecasting methods such as models, gradient boosted decision trees, and neural networks demonstrated considerable effectiveness. The research also addressed the incorporation of external data and validation techniques, as well as how the characteristics of the data influence the selection of statistical or machine learning approaches. Nonetheless, there was no explicit examination of the strengths and weaknesses of the chosen Kaggle dataset [8].

2. Literature Review

Several studies have been conducted on spam detection on social media, specifically on the Twitter platform including using machine learning-based approach that uses lexical, syntactic, and semantic features, along with decision trees, naive Bayes, and support vector machines as classifiers. For instance, a paper explored various spam filtering solutions, including Bayesian classifiers and support vector machines, and found that SVM and Naive Bayes outperformed other models in terms of spam identification Thomas and Meshram introduced a Deep Neuro Fuzzy Network driven by Chimp

Sailfish Optimization (ChSO-based DNFN) for screening out spam content on social media platforms. This method exhibited enhanced performance with regards to precision, recall, and F-measure [9]. Another study investigated on Twitter spam accounts detection and found that ensemble and majority voting machine learning models improved prediction accuracy compared to individual models, with Random Forest being the best model for detecting Twitter spam accounts [10].

Machine learning techniques have been used for spam detection on Twitter. Several studies have explored the use of machine learning models like Decision Tree (DT), Support Vector Machine (SVM), and Naïve Bayesian (NB), Logistic Regression (LR), Random Forest, K-Means, and Agglomerative clustering for this purpose. These models have been applied to datasets obtained from the Twitter API and publicly accessible datasets. The results have shown that the Naïve Bayesian classifier achieved the highest accuracy of 96.83% in detecting spam accounts on Twitter In addition, ensemble models and the majority voting approach have shown to enhance the accuracy in predicting Twitter spam account detection when compared to individual models. [11]. Furthermore, a fusion of content-based features has been applied in conjunction with machine learning techniques like Naive Bayes, Logistic Regression, and K-Nearest Neighbors (KNN), Decision Tree, and Support Vector Machine (SVM) for the purpose of discerning spam tweets on Twitter. [12].

Multinomial Naïve Bayes, with other state-of-the-art models and found that it achieved the highest accuracy for spam detection on social media Past research has pinpointed various recurrent difficulties encountered in detecting spam on social media. These obstacles encompass the divisive sentiments sparked by social spam, its influence on user engagement duration and the quality of available information, as well as the utilization of synchronized automated accounts or bots to disseminate spam. Additional hurdles involve the rise of Deepfakes powered by artificial intelligence, the resilience of detection systems, the ability to scale, the availability of real-time datasets, and the threat of adversarial attacks targeting machine learning-driven spam detectors. To address these challenge, previous studies have employed various techniques such as dimensionality reduction, feature selection/extraction, and machine learning and deep learning algorithms for social spam and spammer detection. They have also explored countermeasures for Deepfake spam and discussed emerging issues and possible directions for future research [13].

Scholars have discovered that employing ensemble models and employing a majority voting approach can enhance the accuracy of predicting Twitter spam account detection when compared to individual models. [14]. Additionally, a Deep Neuro Fuzzy Network driven by Chimp Sailfish Optimization (ChSO-based DNFN) has been suggested as an efficient technique for screening out spam content on various social media platforms, encompassing Twitter. Another study compared the performance of different machine learning classifiers like decision trees and support vector machines, naive Bayesian, and logistic regression, and found that the naive Bayesian classifier provided the best results in terms of accuracy for distinguishing between spammer and authentic account tweets on Twitter [15].

Recent trends and developments in spam detection research on the Twitter platform utilized deep learning algorithms like CNN. In the realm of spam detection, the utilization of CNN has been employed. Enhancing the effectiveness and precision of detection techniques is a paramount objective. In one study, it is suggested to integrate a CNN with an attention model as a means of detecting network spam, which automatically extracts features and achieves fast and efficient detection Another paper uses a deep learning model with a Deep Residual Convolutional Neural Network (DRCNN) to identify potential multiple attacks and intrusions in the network, achieving high classification accuracy. These investigations showcase the proficiency of CNN spam and recognizing diverse network attacks [16].

Some research endeavors have undertaken comparisons between the efficacy of CNNs and alternative machine learning techniques in detecting spam on various social media platforms [17]. Another paper analyzed prior research on spam message identification in social networking platforms and explained the process of detecting spam messages using machine learning models. Another study proposed an effective machine learning-based approach for spam detection on social media, using features like lexical, syntactic, and semantic feature, and classifiers like decision trees, naive Bayes, and support vector machines [18]. Twitter spam account detection found that ensemble and majority voting models outperformed individual models [19].

Preprocessing data is important in the context of spam detection on social media. Conventional methods for word representation based on frequency are both time-intensive and less effective in generating contextual word vectors. To ensure precise information and a comprehensive grasp of the entire procedure, an examination of previous studies pertaining to the detection of spam communications within social networking platforms was conducted [20]. Conducting preprocessing using NLP and subsequently implementing the SVM CTM approach on the refined data for spam analysis has shown a better prediction ratio. There are specific approaches and techniques implemented to address class imbalance in spam detection datasets on Twitter. An option involves employing The Deep Neuro Fuzzy Network powered by Chimp Sailfish Optimization (ChSO-based DNFN). This method effectively filters spam information in social media by utilizing deep learning classifiers and performs well under high dimensional data [21].

The implementation of deep learning models on spam detection in social media can have a significant impact on user experience and platform security. By accurately identifying and filtering out spam messages, machine-learning models can ensure that users receive relevant and legitimate content, enhancing their overall experience on the platform. Additionally, the detection of spam messages helps to protect users from potentially harmful or fraudulent information, safeguarding their security. Diverse learning techniques, including CNN and RNN have been utilized for this purpose. enhance the precision of spam detection. These models have exhibited encouraging outcomes in identifying spam messages and are adept at managing extensive dataset [22]

Spam detection on Twitter presents specific challenges compared to other social media platforms. These challenges include the limited length of tweets, the use of slang words, symbols, and abbreviations, and the rapid spread of spam due to the platform large user base. Previous research has addressed these challenges by employing various techniques. Studies have been conducted on data security and privacy in the context of spam detection on social media, particularly on Twitter. To address the issues, various communities have experimented with various learning models, encompassing individual models, ensemble methods, and majority voting models, in order to enhance the precision of predicting Twitter spam detection [23].

3. Proposed Method

3.1 Mathematical Concept

The proposed model is a neural network architecture known as Convolutional Neural Network (CNN), designed for binary classification tasks. Let's explain the mathematical concepts of each layer in the model. The Conv1D layer equipped with 128 filters and a kernel size of 3 initiates the CNN feature extraction process. In this layer, 128 different filters are applied to the input data, each designed to detect specific patterns or features within a local receptive field of size 3. The convolution operation involves taking the dot product of the filter weights and the corresponding input values, followed by an element-wise application of the Rectified Linear Unit (ReLU) activation function. This function introduces non-linearity by setting negative values to zero, allowing the model to learn complex relationships and abstract representations in the data. We utilized one-dimensional features in this research, 1D-CNN conducts convolutional operations on the local input signal, and various kernels generate distinct input signals. Each grain detects distinct qualities in any location on the input features book. Equation 1 describes the formula of 1D CNN [25], and Table 1 represents the equation symbols.

$$x_j^l = f\left(\sum_{i=1}^{M} x_i^{l-1} . k_{ij}^l + b_j^i\right) \quad (1)$$

Table 1. Equation description

Symbol	Describe
$\mathbf{l_i}$	1D Convolutional layer
$\boldsymbol{k}$	Number of convolutional kernels
$\boldsymbol{j}$	Kernel size
$\boldsymbol{M}$	Channel input number
$\boldsymbol{b}$	Kernel bias
$\boldsymbol{f}$	Activation function

To improve the accuracy of training and testing, a study can utilize several parameters in the DL algorithm, including epoch, regular, and optimizer. Figure 1 illustrates the CNN topology for training the book lending recommendation model.

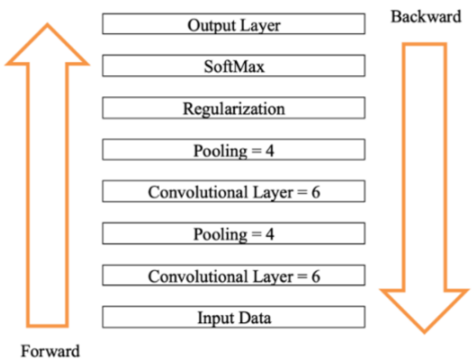

Figure 1. The proposed CNN topology for training the book lending dataset.

The proposed CNN employs a Deep Neural Network with several hidden layers to train and test the model. It also utilizes a gradient descent to minimize the objective function with the model's parameters. The model updates the settings in the opposite direction of the gradient of the objective function. Different from the regular pooling layer, this study adopts a pooling layer to optimize and accelerate training time in a neural network. By using the proposed CNN topology, we calculate the accuracy and loss of the training and testing process to achieve the best result with the diverse input vector. This study calculates the 1D dataset of OSN, so tuning an appropriate hyperparameter is beneficial [23][24][25]. We establish a supervised learning model by defining calculation over NN as follows:

Input features $x^{(i)} \in R$
Outputs $x^{(i)} \in Y\ (e.g. R, \{0, 1\}, \{1, \ldots, p\})$
Model parameters $\theta \in \mathbb{R}^k$
Hypothesis function $h_{\theta:} : \mathbb{R}^n \rightarrow \mathbb{R}$
Loss function $\ell : \mathbb{R} \times Y \rightarrow \mathbb{R}_{+:}$

On a CNN we need to calculate forward pass and backward pass to measure the gradient of the loss function in the model. The study calculates the forward pass to convolve input matrix x_i with filter W_i to produce convolution output $z_{i:}$ as follows:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m \tag{2}$$

$$z_{i:}(x_i) = W_i\, x_i + b \tag{3}$$

3.2 Datasets

This study utilized dataset of Twitter spam from the NSCLab. It comprises a total of 10,001 records, each encompassing 13 distinct features. The target variable for this dataset is binary, categorized as either 'yes' or 'no', indicating the presence or absence of spam content. This comprehensive dataset serves as the foundation for investigating and developing methods to effectively discern spam-related activities within the Twitter platform.

3.3 Preprocessing

In this study, a series of preprocessing steps were employed to prepare the dataset for analysis. Firstly, irrelevant features were discarded to streamline the data. Next, any rows containing missing values (NAN) were removed to ensure data integrity. Subsequently, categorical string data underwent conversion into numerical format, facilitating compatibility with machine learning algorithms. Additionally, the age feature underwent a normalization process to bring it within a standardized range. Finally, to evaluate the model performance, The dataset underwent a partitioning process, where 80% of the data was designated for training purposes, while the remaining 20% was set aside for testing. These preprocessing steps collectively ensured that the dataset was appropriately refined and formatted for subsequent analysis and model development.

4. Result and Analysis

In the experiment result, Fig. 1 portrays the training accuracy (blue line) and validation accuracy (orange line), providing fundamental insights into the model's learning progression. By scrutinizing trends over epochs, potential issues like overfitting or underfitting can be discerned. Additionally, it serves as an initial assessment of the model's ability to generalize to unseen data. Fig. 2, on the other hand, depicts the training loss (blue line) and validation loss (orange line), shedding light on the model's convergence and optimization throughout training. Observing trends in loss over epochs aids in understanding the model's error minimization and generalization capabilities. Discrepancies between training and validation loss curves can signal overfitting or underfitting.

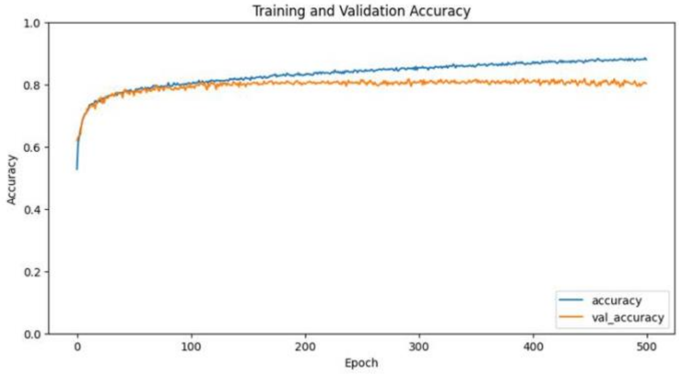

Fig 1. Training and Validation Accuracy.

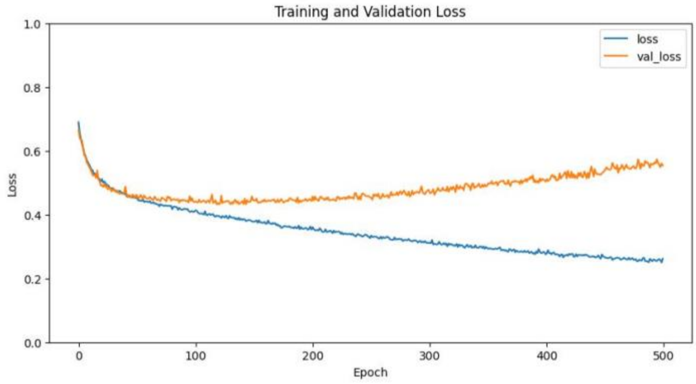

Fig. 2. Training and Validation Loss.

4.2. Model Performance

In this part, Table 1 and Table 2 presented a comprehensive evaluation of the algorithms utilized in this study: CNN, SVM, Decision Tree, KNN, Gaussian Naive Bayes (GNB), and Gradient Boosting (GBoost). Table 1 offers a detailed Confusion Matrix for each algorithm, delineating the counts of True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN). This granular insight enables a thorough assessment of their respective abilities in class distinction, contributing to a nuanced understanding of their effectiveness within the study's context. Meanwhile, Table 2 encapsulates the Classification Report, furnishing critical performance metrics including accuracy, precision, recall, and F1-score for each algorithm. Accuracy denotes the proportion of correctly classified instances, while precision evaluates positive.

Table 1. Confusion Matrix

Algorithm	TP	FP	FN	TN
CNN	797	208	154	841
SVM	604	401	182	813
Decision Tree	819	186	297	698
KNN	736	269	251	744
GNB	181	824	52	943
Gboost	819	186	193	802

Table 2. Classification Report

Algorithm	Accuracy	Class	Precision	Recall	F1- score
CNN	81.89	HAM	0.84	0.79	0.81
		SPAM	0.80	0.85	0.82
SVM	70.51	HAM	0.77	0.60	0.67
		SPAM	0.67	0.82	0.74
DT	48.66	HAM	0.73	0.81	0.77
		SPAM	0.79	0.70	0.74
KNN	74.00	HAM	0.75	0.73	0.74
		SPAM	0.73	0.75	0.74
GNB	48.66	HAM	0.78	0.18	0.29
		SPAM	0.53	0.95	0.68
Gboost	81.05	HAM	0.81	0.81	0.81
		SPAM	0.81	0.81	0.81

6. Conclusion

The study underscores the outstanding effectiveness of CNNs in detecting Twitter spam, surpassing traditional algorithms like SVM, Decision Tree, KNN, Gaussian Naive Bayes, and Gradient Boosting. The CNN harvested superior performance in accuracy, precision, recall, and F1-score, demonstrating its exceptional capacity to differentiate between spam tweets and genuine content. While constraints such as dataset size and assumptions in data preprocessing may constrain the generalizability of the results, the thorough research methodology ensures the credibility of the findings within the specified scope. This research not only tackles a crucial issue in online security but also contributes to the continual progress of machine learning methods for textual analysis and classification. Furthermore, practical suggestions for real-time spam detection systems and user-driven spam flagging, coupled with future research directions like multimodal spam identification and analyzing spammers' behaviors, present promising avenues for enhancing spam detection and reinforcing online security measures. Ultimately, this study offers valuable insights and recommendations with substantial implications for the broader landscape of online security and spam identification.

Acknowledgment

This paper is conducted in the Department of Informatics, Saba University, Yemen.

Data Availability

The dataset comes from the Kaggle website (https://www.kaggle.com/datasets/andrewmvd/fetal-health-classification).

References

[1] M. Thomas and B. B. Meshram, "ChSO-DNFNet: Spam detection in Twitter using feature fusion and optimized Deep Neuro Fuzzy Network," Advances in Engineering Software, vol. 175, p. 103333, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103333.

[2] L. Das, L. Ahuja, and A. Pandey, "Analysis of Twitter Spam Detection Using Machine Learning Approach," in 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), Apr. 2022, pp. 764–769. doi: 10.1109/ICIEM54221.2022.9853100.

[3] S. J. Alsunaidi, R. T. Alraddadi, and H. Aljamaan, "Twitter Spam Accounts Detection Using Machine Learning Models," in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Dec. 2022, pp. 525–531. doi: 10.1109/CICN56167.2022.10008339

[4] V. Niranjani, Y. Agalya, K. Charunandhini, K. Gayathri, and R. Gayathri, "Spam Detection for Social Media Networks Using Machine Learning," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), Mar. 2022, pp. 2082–2088. doi: 10.1109/ICACCS54159.2022.9785149.

[5] R. M. Gomathi, M. R. Reddy, M. V. Bhadra Reddy, T. Anandhi, A. Sivasangari, and P. Ajitha, "Spam Review Detection using Linguistic Methods for Specified user in Twitter," in 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), May 2022, pp. 1743–1748. doi: 10.1109/ICICCS53718.2022.9788453.

[6] A. M. Soomro, A. B. Naeem, S. Bagchi, N. Sharma, P. Singh, and S. K. Debnath, "Uncovering Spam in Twitter: A Machine Learning Approach," in 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Apr. 2023, pp. 993–998. doi: 10.1109/CISES58720.2023.10183473.

[7] S. J. Alsunaidi, R. T. Alraddadi, and H. Aljamaan, "Twitter Spam Accounts Detection Using Machine Learning Models," in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Dec. 2022, pp. 525–531. doi: 10.1109/CICN56167.2022.10008339.

[8] R. G. Jimoh et al., "Experimental Evaluation of Ensemble Learning-Based Models for Twitter Spam Classification," in 2022 5th Information Technology for Education and Development (ITED), Nov. 2022, pp. 1–8. doi: 10.1109/ITED56637.2022.10051587.

[9] M. Thomas and B. B. Meshram, "ChSO-DNFNet: Spam detection in Twitter using feature fusion and optimized Deep Neuro Fuzzy Network,"Advences in Engneering Sofware, vol. 175 p. 103333, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103333.

[10] A. Srivastava, B. Jha, M. S. Fahad, A. Deepak, and K. Abhishek, "A Genetic Algorithm Formulation for Rogue Taxa Problem," in 2018 International Conference on Bioinformatics and Systems Biology (BSB), Oct. 2018, pp. 161– 164. doi: 10.1109/BSB.2018.8770676.

[11] D. Łydżba, A. Różański, and D. Stefaniuk, "Equivalent microstructure problem: Mathematical formulation and numerical solution," International Journal of Engineering Science, vol. 123, pp. 20–35, Feb. 2018, doi: 10.1016/j.ijengsci.2017.11.007.

[12] B. Wenkai, "Study on Vehicle Recognition Algorithm Based on Convolutional Neural Network," in 2023 3rd Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS), Feb. 2023, pp. 212–216. doi: 10.1109/ACCTCS58815.2023.00083.

[13] H. A. Ibrahim, M. N. Thamilvanan, A. Zaian, and E. Supriyanto, "Fertility Assessment Model For Embryo Grading Using Convolutional Neural Networ (CNN)," in 2022 International Conference on Healthcare Engineering (ICHE), Sep. 2022, pp. 1–4. doi: 10.1109/ICHE55634.2022.10179864.

[14] N. W. Parwati Septiani et al., "Convolutional Neural Network (CNN) Algorithm for Geometrical Batik Sade' Motifs," in 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Feb. 2023, pp. 597–602. doi: 10.1109/ICCoSITE57641.2023.10127829.

[15] M. Akbari and J. Liang, "Semi-Recurrent Cnn-Based Vae-Gan for Sequential Data Generation," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr. 2018, pp. 2321–2325. doi: 10.1109/ICASSP.2018.8461724.

[16] M. Abulaish and A. K. Sah, "A Text Data Augmentation Approach for Improving the Performance of CNN," in 2019 11th International Conference on Communication Systems & Networks (COMSNETS), Jan. 2019, pp. 625–630. doi: 10.1109/COMSNETS.2019.8711054.

[17] P. Shivakumara, A. Banerjee, U. Pal, L. Nandanwar, T. Lu, and C. -L. Liu, "A New Language-Independent Deep CNN for Scene Text Detection and Style Transfer in Social Media Images," IEEE Transactions on Image Processing, vol. 32, pp. 3552–3566, 2023, doi: 10.1109/TIP.2023.3287038.

[18] M. Akbari and J. Liang, "Semi-Recurrent Cnn-Based Vae-Gan for Sequential Data Generation," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr. 2018, pp. 2321–2325. doi: 10.1109/ICASSP.2018.8461724.

[19] S. J. Alsunaidi, R. T. Alraddadi, and H. Aljamaan, “Twitter Spam Accounts Detection Using Machine Learning Models,” in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Dec. 2022, pp. 525–531. doi: 10.1109/CICN56167.2022.10008339.
[20] M. Thomas and B. B. Meshram, “ChSO-DNFNet: Spam detection in Twitter using feature fusion and optimized Deep Neuro Fuzzy Network,” Advances I Engineering Software, vol. 175, p. 103333, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103333.
[21] S. Kishore, M. Awad, and A. Al-Zubidy, “Spam Detection Techniques Recapped,” in 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Nov. 2022, pp. 157–162. doi: 10.1109/ICECTA57148.2022.9990450.
[22] S. Kishore, M. Awad, and A. Al-Zubidy, “Spam Detection Techniques Recapped,” in 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Nov. 2022, pp. 157–162. doi: 10.1109/ICECTA57148.2022.9990450.
[23] C. S. Bojer and J. P. Meldgaard, “Kaggle forecasting competitions: An overlooked learning opportunity,” International Journal of Forecasting, vol. 37, no. 2, pp. 587–603, Apr. 2021, doi: 10.1016/j.ijforecast.2020.07.007.
[24] M. Thomas and B. B. Meshram, “ChSO-DNFNet: Spam detection in Twitter using feature fusion and optimized Deep Neuro Fuzzy Network,” Advances in Engineering Software, vol. 175, p. 103333, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103333.
[25] S. J. Alsunaidi, R. T. Alraddadi, and H. Aljamaan, “Twitter Spam Accounts Detection Using Machine Learning Models,” in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), Dec. 2022, pp. 525–531. doi: 10.1109/CICN56167.2022.10008339.