Hindawi
Wireless Communications and Mobile Computing
Volume 2022, Article ID 2215852, 11 pages
https://doi.org/10.1155/2022/2215852

Research Article

A Deep Learning-Based Framework for Feature Extraction and Classification of Intrusion Detection in Networks

Muhammad Naveed,[1] **Fahim Arif**,[2] **Syed Muhammad Usman**,[3] **Aamir Anwar**,[4] **Myriam Hadjouni**,[5] **Hela Elmannai**,[6] **Saddam Hussain**,[7] **Syed Sajid Ullah**,[8] **and Fazlullah Umar**[9]

[1] *Department of Computer Science, SZABIST, Islamabad, Pakistan*
[2] *Department of Computer Software Engineering, MCS, NUST, Islamabad, Pakistan*
[3] *Department of Creative Technologies, Air University, Islamabad, Pakistan*
[4] *School of Computing and Engineering, The University of West London, UK*
[5] *Department of Computer Sciences, College of Computer and Information Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia*
[6] *Department of Information Technology, College of Computer and Information Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia*
[7] *School of Digital Science, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam*
[8] *Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway*
[9] *Department of Information Technology, Khana-e-Noor University, Pol-e-Mahmood Khan, Shashdarak, 1001 Kabul, Afghanistan*

Correspondence should be addressed to Saddam Hussain; saddam_1993@hotmail.com,
Syed Sajid Ullah; sajidullah718@gmail.com, and Fazlullah Umar; fazlullahumer@gmail.com

Received 28 April 2022; Accepted 1 July 2022; Published 8 August 2022

Academic Editor: Kuruva Lakshmanna

An intrusion detection system, often known as an IDS, is extremely important for preventing attacks on a network, violating network policies, and gaining unauthorized access to a network. The effectiveness of IDS is highly dependent on data preprocessing techniques and classification models used to enhance accuracy and reduce model training and testing time. For the purpose of anomaly identification, researchers have developed several machine learning and deep learning-based algorithms; nonetheless, accurate anomaly detection with low test and train times remains a challenge. Using a hybrid feature selection approach and a deep neural network- (DNN-) based classifier, the authors of this research suggest an enhanced intrusion detection system (IDS). In order to construct a subset of reduced and optimal features that may be used for classification, a hybrid feature selection model that consists of three methods, namely, chi square, ANOVA, and principal component analysis (PCA), is applied. These methods are referred to as "the big three." On the NSL-KDD dataset, the suggested model receives training and is then evaluated. The proposed method was successful in achieving the following results: a reduction of input data by 40%, an average accuracy of 99.73%, a precision score of 99.75%, an F1 score of 99.72%, and an average training and testing time of 138% and 2.7 seconds, respectively. The findings of the experiments demonstrate that the proposed model is superior to the performance of the other comparison approaches.

1. Introduction

There has been a discernible increase in the volume of traffic on the network. On the other hand, the number of potential infiltration threats has grown and their level of sophistication has also improved. Communication that is reliant on networks is now susceptible to attacks from both the outside and the inside. It is quite difficult to check incoming traffic since there is a large volume of traffic and a high number of attacks, which also increases the amount of time and money spent on computing. For this purpose, researchers are motivated to design an intelligent detection system that uses less computational time than traditional methods but gives a high level of accuracy.

An IDS is widely used for the classification of network traffic to identify anomalies inside the network. IDS is the software that analyses real-time network traffic and reports any abnormal activity going over the network. IDS can be divided into two types of systems: signature and anomaly-based detection systems. Signature-based IDS uses predefined patterns and matches the incoming traffic with existing patterns. If a match is not found, it classifies it as an anomaly, otherwise a normal pattern. The signature-based method cannot detect unknown and new attacks, whereas the anomaly-based technique is intelligent enough to identify any unknown attack on a network. Researchers prefer anomaly-based intrusion systems to handle unknown and unauthorized access on the network. However, anomaly-based systems give low accuracy and a high-false-alarm rate while dealing with high-dimensional data [1]. Researchers have also proposed a hybrid approach that combines both signature and anomaly-based approaches to handle seen and unseen data. In the hybrid approach, the computational cost is very high and the system gives poor performance in terms of accuracy [2–4].

Multiple machine/deep learning methods [5–15] for detecting intrusion in networks have been proposed in recent years; however, data dimensionality remained one of the biggest problems in intrusion systems. Due to high-dimensional data, IDS suffers in performance and accuracy. One of the solutions to this issue is to cut down on the amount of input features and make use of only those features that are reliable and have a significant bearing on the category of the final result. The purpose of feature selection is to select an optimal feature subset less than the original dataset and provide an efficient system with better accuracy. In network classification, data can contain some irrelevant features that can increase system computational time and affect accuracy. Feature selection techniques help us remove irrelevant data. Feature selection is considered a vital step in preprocessing as it can affect the system performance if relevant features are not removed from the original dataset [16, 17].

Feature selection algorithms are categorized into filter-based and wrapper-based techniques. Wrapper methods provide the best relevant feature subset, but they cost more computation time, which degrades the system performance. Similarly, filter methods are computationally efficient, have fast processing speeds, and are less prone to overfitting [18]. With the rapid increase in network traffic, intrusion detection systems are facing data dimensionality and system complexity issues. Feature selection is becoming an important phase of preprocessing for network classification problems. Feature selection helps us to reduce and remove irrelevant and redundant features from the main dataset that have no impact on classification results. The feature selection method selects a subset from the original dataset using some criteria that contain the properties of the original dataset. According to Kantardzic [19], when features are reduced from large datasets using basic techniques, classification improves. The parameters are discussed as follows:

1.1. Less Computational Power. When a large dataset is reduced using feature selection techniques, it also reduces system computational power as less time is required to train and test the model on the reduced dataset.

1.2. Improved Detection Accuracy. During the feature selection process only, those features are removed which have very low or no impact on classification so removing noisy features helps improve model accuracy. There are two main techniques of the feature selection wrapper method and the filter method. Both techniques have different advantages and disadvantage as described in Table 1.

In this research paper, we propose a two-stage hybrid model. In the first stage, we applied filter-based feature selection techniques to reduce the dimensionality of input data. After getting the optimal feature subset, we have used the deep learning model (DNN) [20–22] for classification and have achieved increased accuracy with less processing time.

2. Related Work

In this section, we will take a look at some of the most recent accomplishments that have been made in the field of anomaly and intrusion detection. IDS is an essential component of a secure network because it monitors the traffic that occurs between all of the devices connected to the network. There has been a significant amount of study conducted in the academic literature on the subject of identifying anomalous patterns of behavior, and numerous machine learning, deep learning, and hybrid approaches have been employed [23]. An IDS is a type of security management system that monitors the traffic coming into and going out of a computer system in order to identify any harmful behavior that may be taking place over a network. These systems examine the information coming from all of the sources before sending it on to the network for further processing. There are multiple features used by these systems to detect intrusion. Intrusion detection and protection systems are divided into four major categories: network based, wireless based, network behavior analysis based, and host based [24].

The usage of deep belief networks, often known as DBN, is common in IDS. The DBN has the power to learn high-dimensional representations of data in addition to doing categorization in an effective and precise manner. In order to fine tune the DBN model for improved classification, only a very little amount of labelled data is required [25]. On the KDD 99 dataset, the performance of the DBN is evaluated and it demonstrates superiority to both the SVM and ANN classification models that are currently in use. Potluri and Diedrich came up with the idea for a DNN-based intrusion detection system that can classify attacks. According to the findings, the suggested model is more successful at identifying classes of DoS and probe objects but it is less successful at identifying classes of R2L and U2R. Because there was little data available for training purposes, the detection accuracies were inconsistent in R2L and U2R cases but were

Table 1: Advantages and disadvantages of feature selection techniques.

Selection techniques	Advantages	Disadvantages	Methods
Filter	(i) Computationally efficient (ii) Higher processing power (iii) Independent of classification algorithm	Can select irrelevant features	(i) ANOVA (ii) Chi square (iii) Pearson correlation
Wrapper	(i) Highly accurate feature subset	High computational cost; requiring leaning algorithm	(i) Recursive technique (ii) Forward selection (iii) Backward elimination

reliable in DoS and probe situations [26, 27]. Kim et al. [27] proposed an intrusion detection system that was based on deep neural networks (DNN). In hidden layers, the activation function that is used is called ReLU.

A lightweight deep learning model was proposed by Zeng et al. [28], which makes use of deep learning for the classification of encrypted traffic and the detection of intrusions. Due to the deep learning usage model, they were able to understand unseen traffic. Results prove that the proposed model is more reliable and accurate with a minimum use of resources. Similarly, in [29], a ML- and DL-based technique is proposed. As technology improves, the number of threats to networks is always changing. Because of this, not all public datasets have data on all types of threats and attacks. Due to the dynamic nature of attacks, models underperform against unseen and unpredicted data. Due to the unseen problem of models, a new approach is proposed which basically classifies the unseen and unpredictable attacks. The model is trained on the latest datasets containing almost all types of cyberattacks, which makes it highly scalable and hybrid in the DNN framework.

Intrusion detection systems need high accuracy and detection time to compete with modern cyberattacks. A scale-hybrid intrusion detection and alert system was presented by Vinayakumar et al. [30]. The framework enables real-time monitoring of network traffic and the notification of system administrators of potentially harmful activity on the network. It was stated that the system would provide a DNN architecture that is both effective and heterogeneous and that it would be able to manage and analyse huge volumes of data in real time. Several other datasets, such as NSL-KDD and KDD'99, were utilised in the evaluation of the architecture. The best F-measure for binary classification on NSL-KDD was 80.7%, and the best F-measure for multiclass classification was 76.5% [30, 31].

A DNN-based model for anomaly detection in software-defined networks has been proposed by Tang et al. [32]. The proposed model has one input layer, three hidden layers, and one output layer. All of these layers are concealed from view. The NSL-KDD dataset served as the basis for some experiments. Only six out of the total of forty-one features are put to any kind of practical use, and the subset of these six features came from an SDN environment. When applied to a binary classification task, the model demonstrated an accuracy of 75.75%. The BAT model for the intrusion detection system was proposed by Su et al. [33]. The bidirectional long-short-term memory (BLSTM) and attention mechanism are the two components that make up the BAT model. The model shows better accuracy on the NSL-KDD test dataset and requires 100 epochs to be trained. In multiclass problem, the BAT model shoes 3% and 4% higher accuracy than CNN and RNN, respectively. The few-shot learning-based method is presented by Yu & Bian [34] to increase the network-based security and allow efficient intrusion detection. The proposed model achieves highest accuracy of 92.34% in detecting abnormal network behaviors, and the model is evaluated on NSL-KDD and UNSW-NB15 datasets. The model is trained using 2% data and still achieves leading performance.

Ahmadi et al. [35] have proposed a hybrid approach to improve the efficiency of IDS. The subset received from these techniques is passed to the decision tree classifier for classification results. The proposed feature selection model returns 20 useful features out of 41 features of the NSL-KDD dataset. The highest accuracy achieved by the classification model is 80.6. A novel feature selection and classification model is proposed by Ahmadi et al. [36]. The feature selection model uses chi square, information gains, and correlation-based techniques which are used with majority voting. The majority voting model return optimal features which are passed to the decision tree for classification purposes. The proposed model achieved around 80% accuracy whereas a total of 20 features are used from the NSL-KDD dataset. Similarly, the GAN-based feature selection and oversample handling scheme is proposed [37]. Dimensionality reduction and oversampling are one of the core issues in classification problem especially intrusion detection systems. Results show that the proposed model returns better features and enhances the classification model's performance. Feature selection is considered one of the main parts of IDS because these systems have to deal with a large amount of data so a strong feature reduction technique is always encouraged to be applied with the network classification problem. Researchers [34, 36–52] have used different feature selection techniques. The gain ratio, Pearson correlation, and ANOVA are few of the techniques that are widely used. Feature selection helps us to reduce the input data size by removing redundant and irrelevant features and features with no impact on classification [53–62].

3. Methodology

In this section, the overall methodology of the article will be presented. The methodology of the paper is divided into two stages. The first phase is known as data preprocessing, and it includes processes such as data

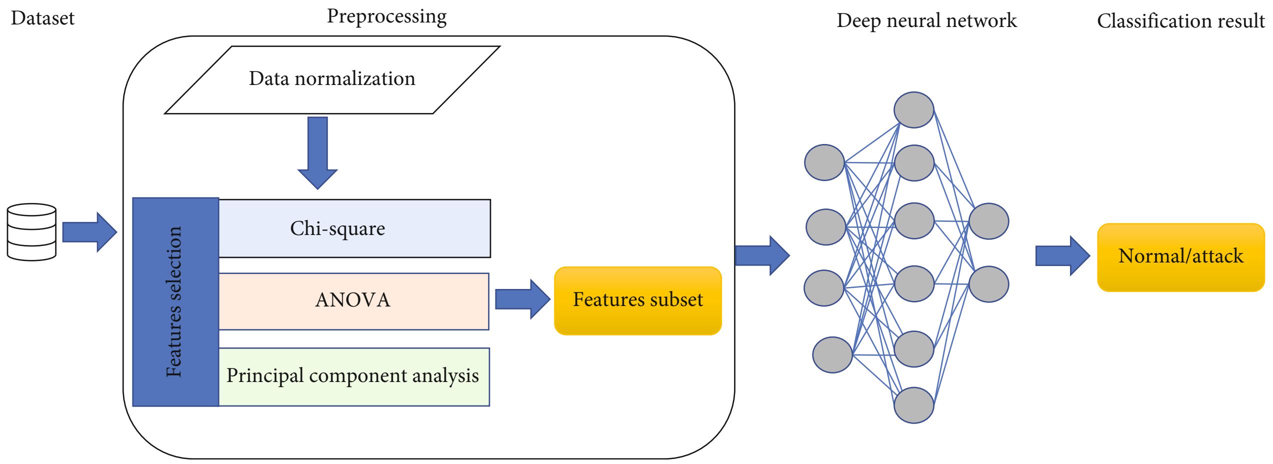

Figure 1: Block diagram of the proposed model.

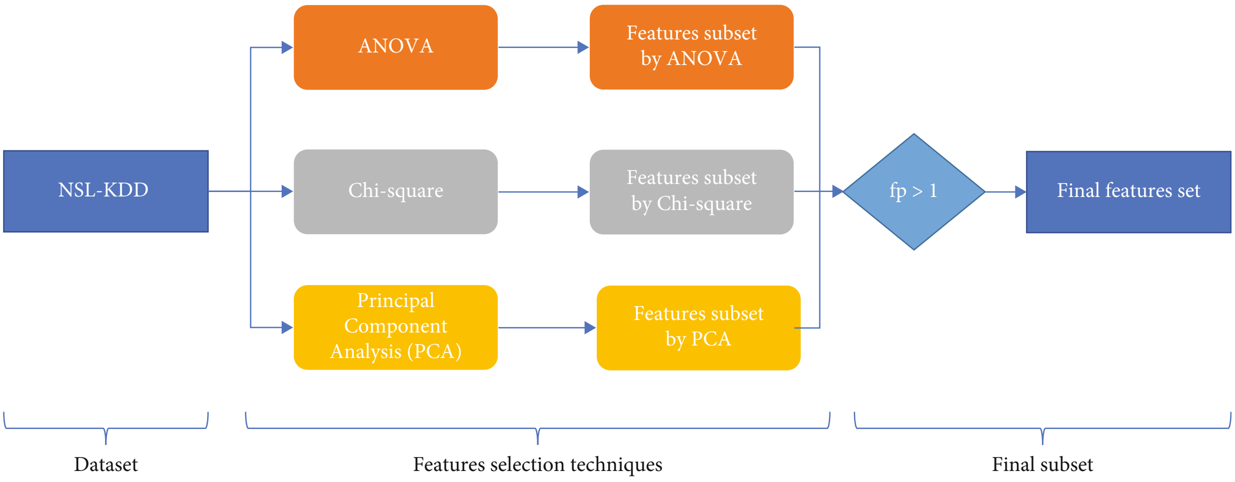

Figure 2: Proposed feature selection model.

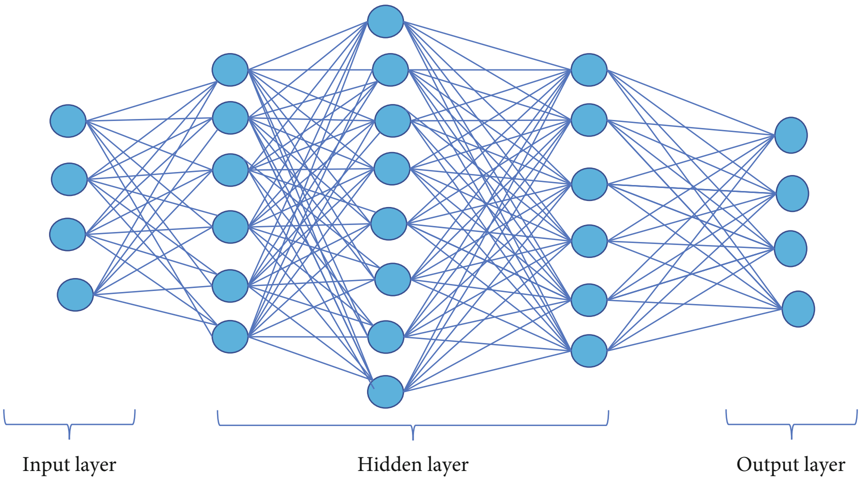

Figure 3: General DNN model.

normalization, data encoding, and feature selection. The second phase is known as the deep neural network model, and it is responsible for getting the preprocessed data and classifying the traffic as either normal or abnormal. The block schematic of the suggested model can be seen in Figure 1.

```
Input: NSL-KDDTrain++2    Output: Accuracy, Precision, Recall, F1- Score;
Initialization:
 f = features, nfeatures = Numeric features, tfeatures= textual features, f_c =
 features from chi-square, f_a= features from ANOVA, f_p= Features from,
 PCA, f_n= Final Features subset, x= number of times a feature repeat in any
 three subsets (f_c, f_a, f_p)
Step 1: Data Preprocessing
    f=MinMaxnormalization(f)
    nfeatures'=encodenumericzscor(nfeatures)
    tfeatures'=encodetextdummy(tfeatures)
 EndStep
Step 2: Features Selection
    f_c=Chi-Square(f)
    f_a=ANOVA(f)
    f_p =PCA(f)
EndStep
Step 3: Classification
    Model is trained and tested on NSLKDD Binary Classification Dataset.
    Relu is used in input and Hidden layers while Sigmoid in Output layer
    EndStep
    Return the classification result.
```

Algorithm 1: A deep learning-based framework for feature extraction and classification for intrusion detection in networks.

3.1. Data Preprocessing. The purpose of data preprocessing is to optimize the information collection and processing by making adjustments to the values of the data in a particular dataset. Because there is usually a significant difference between the dataset's maximum and minimum values, normalizing the data reduces the algorithm's complexity. According to Chiba et al., the results of classification can be improved with proper data preprocessing specially in deep learning [63].

3.1.1. Data Normalization. The NSL-KDD dataset contains both discrete and continuous features, the same as KDD99 [64]. Difference in feature values makes features more diverse and contrasting. So, the preprocessing phase is required to normalize the data and scale all feature values into the same range. Features are normalized using mean and standard deviation to make the same value range. Equation (1) describes the mean algorithm used for feature scaling.

$$\text{Mean} = \frac{1}{t \times \sum_{k=1}^{n} (xk)}. \tag{1}$$

Here, the mean is an arithmetic mean. T is the total no. of rows in a single column that are being averaged. X_k is the individual averaged value; we use standard deviation to handle the data dispersion. The dataset contains multiple features with widespread values for which deviation is required. The formula of standard deviation used in paper is given in equation (2).

$$\text{SD} = \sqrt{\frac{\sum_1 (x_i - \bar{x})^2}{N - 1}}, \tag{2}$$

where x_i is the i^{th} point in the dataset, $\bar{x}$ is the mean value of the dataset, and n is the total data points in the dataset.

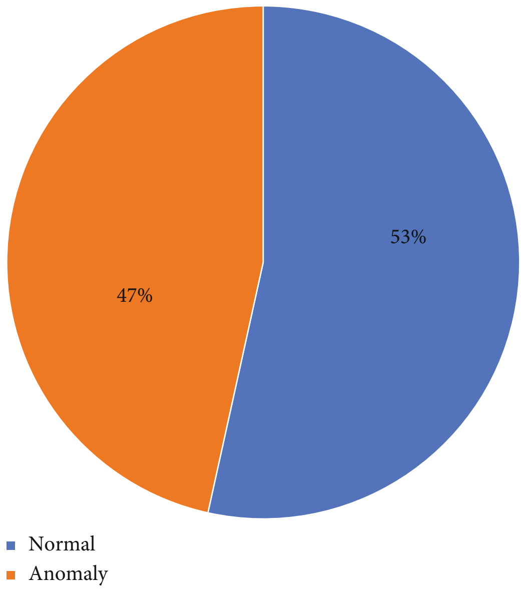

Figure 4: NSL-KDD binary class labels.

Table 2: Feature selection results.

Technique name	No. of features selected
Chi square	26
ANOVA	29
Principal component analysis	25
Final features selected	27

3.2. Feature Selection Techniques. A combination of three filter-based feature selection techniques is used in the feature selection model. The most relevant features were ranked and

Table 3: Comparison of different feature selection techniques with the proposed feature selection model.

Method	Feature selection	Dataset	Features used	Accuracy
Ahmadi et al. [36]	Chi squared, information gain, correlation-based evaluation	NSL-KDD	20	80.6
Liu et al. [37]	ANOVA	NSL-KDD, UNSWNB15, CICIDS-2017	16, 13, 39	83.28
Gottwalt et al. [39]	CorrCorr	NSL-KDD, UNSWNB15	19	95
Vinutha et al. [40]	Chi squared, information gain, gain ratio, correlation-based attribute evaluation, symmetrical uncertainty	NSL-KDD	31	85.91
Tang et al. [41]	SDN environment-based six basic features	NSL-KDD	6	75.75
Bhattacharya et al. [42]	Layered wrapper feature selection approach	NSL-KDD	16	83.14
Rama et al. [43]	Hyper graph-based genetic algorithm (HG-GA)	NSL-KDD, KDD-99	35	97.14
Mohammadi et al. [44]	Linear correlation, cuttlefish algorithm, decision tree	KDD99	10	95.03
Gao et al. [45]	CART algorithm	NSL-KDD	17	79.7
Tang et al. [46]	Stacked autoencoder	NSL-KDD	Autoselection	87.74
Proposed model	ANOVA, chi square, PCA	NSL-KDD	27	99.73

used for classification. The most important features that have a strong influence on the output class are prioritized and chosen by the model to classify the network traffic as normal or anomaly. Chi square, ANOVA, and principal component analysis (PCA) are used for feature selection. The results of all three techniques are combined as a single subset with a threshold value more than one. A feature which is repeating in any of the two subsets was used for the final subset. We combined the results of the multiple feature selection technique as it helps to find the most relevant and strong features and improves classification accuracy [38]. Figure 2 describes the complete feature selection model proposed and used in the paper.

Table 4: Proposed model performance for binary classification.

Performance metrics	Accuracy	Precision	Recall	F1 score
Proposed model	99.73	99.75	99.73	99.72

Table 5: Confusion matrix of the proposed model.

	Predicted attack	Predicted normal
Actual attack	33526	88
Actual normal	82	29298

3.2.1. Chi Square. Chi square is a statistical approach widely used for feature selection. It finds the importance of each individual feature with respect to the outcome class. The chi square value is used to determine the dependence of features on the outcome class. In other words, if a feature has a higher chi square value, it is more dependent on the outcome class and is suitable for classification. The mathematical representation of the chi square technique is given in equation (3).

$$X^2 = \sum_{k=1}^{t}\sum_{l=1}^{n}\frac{(M_{k,} - P_{k,})}{P_{k,l}}, \tag{3}$$

where t is the total number of attributes, n is the total number of classes, and $\boldsymbol{M_k}$ and $\boldsymbol{P_k}$ are the actual and predicted values. The higher the value of chi square (X^2), the more the importance of features for the prediction model.

3.2.2. ANOVA (Analysis of Variance). ANOVA is a univariate feature selection technique that ranks the features according to their variance score. The variance score of features determines its impact on the response class. High variance between features of multiple classes reflects that better classification can be done, whereas low variance leads toward poor classification.

3.2.3. Principal Component Analysis (PCA). The PCA is a highly known method used for the reduction of data. PCA utilizes the linear algebra in order to minimize the dimensionality of the data while maintaining its fundamental nature and useful characteristics. Less information is lost when PCA is applied for feature reduction. It is also less sensitive towards noisy data.

3.3. Deep Neural Network Model for Classification. After data cleaning and reduction, the deep neural network is used for classification purposes. The deep neural network is widely used in intrusion and anomaly-based applications. DNN models are divided into input, hidden, and output layers. The DNN optimizes parameters to avoid the classification errors during training time. Complex hidden layer structures make DNN models more accurate and flexible to handle large datasets. Each layer gains a distinct complexity level for all features. The proposed DNN model contains three hidden layers with the rectified linear unit (ReLU) as

Table 6: Comparison of the proposed work with different models.

Model	Dataset	Classifier	AC (%)	Precision	Recall	F1 score
Tang et al. [46]	NSL-KDD	SAAE-DNN	87.74	86.47	84.12	85
Wang et al. [5]	NSL-KDD	RNN	94.19	—	—	—
Al-Qatf et al. [6]	NSL-KDD	STL-IDS	84.96	96.2	76.5	85.2
Ingre et al. [7]	NSL-KDD	ANN	81.2	—	—	—
Tang et al. [8]	NSL-KDD	SDN-DNN	75.75	83	75	74
Yin et al. [9]	NSL-KDD	RNN-IDS	83.28	—	97.09	—
Li et al. [10]	NSL-KDD	GoogLeNet	81.84	81.84	100	90.01
Tama et al. [11]	NSL-KDD UNSW-NB15	TSE-IDS	85.79	88	86.80	87.4
Choudhary et al. [12]	NSL-KDD	DNN	91.7	93.6	92	92.2
Farahnakian et al. [13]	KDD99	DAE	96.53	—	—	—
YU et al. [34]	NSL-KDD UNSW-NB15	CNN	92.33	96.1	95	93
Wang et al. [48]	NSL-KDD	SDAE-ELM1	78.04	95.99	64.12	76.87
Proposed model	NSL-KDD	DNN	99.73	99.75	99.73	99.72

activation function in the hidden layer and sigmoid in the output layer. Our proposed DNN model contains three hidden layers where Adam is used as the optimizer. The general DNN model is shown in Figure 3.

3.4. Algorithm. Proposed algorithm is shown in Algorithm 1. After the input algorithm starts from step 1 where data normalization and feature encodings are done, initially, all feature values range different so it is required to normalize all feature values into a same scale. As the dataset contains both numeric and textual features, so, it is a must to convert them into the same format before we pass them to the classification model. Step 2 of the algorithm is feature reduction where preprocessed features are passed to three different feature selection models which return three different feature sets. From these three feature sets, only those features are shortlisted for the classification model which are selected by any two or all selection models. Feature selection models used in this study are chi square, ANOVA, and PCA. At the end of step 2, we get a single feature subset which is reduced from the original dataset. Step 3 is basically classification; the DNN is used for classification purposes. The final subset of preprocessed, and the selected features are passed to the DNN model for classification.

3.5. Dataset Description. Data from the NSL-KDD dataset is used to develop and test the model under consideration. In anomaly detection, the NSL-KDD dataset is a well-known and benchmark dataset. It is an updated version of the KDD99 dataset. In NSL-KDD, duplicate entries were removed and class imbalance was also improved as compared to that in KDD99 which contains more than 50% duplicate entries due to which its model was overfit most of the time. NSL-KDD containing 41 features with 2 labels (binary classification) is used in our work. The KDDTrain + binary classification dataset from NSL-KDD is used for training and testing purposes. The dataset contains 125974 unique rows. As shown in Figure 4, the dataset contains a balanced binary class.

4. Experimental Results and Discussion

This section presents the proposed model results. Methodology consists of three steps including preprocessing, feature selection and classification, therefore, multiple experimental settings have been used and results are presented by varying feature selection and classification methods. In feature selection, input data is reduced to 40% that helps to improve model performance as the removed features are considered noisy and irrelevant. The classification result also shows better performance than the existing detection models. Results of feature selection and classification are discussed as follows.

4.1. Feature Selection Results. After the data normalization, the next step is to reduce data dimensionality. Three well-known feature selection techniques are used for feature reduction. These techniques return a subset of features from the main dataset. A feature repeating more than once in any of the three subsets was selected for final input for the classification model. Table 2 shows the features generated by all three techniques.

The proposed feature selection technique outperforms the other techniques in terms of accuracy as shown in Table 3. Feature set obtained as a result of our proposed feature selection method gives high classification accuracy, precision and recall with low computational complexity. The model achieved the highest accuracy of 99.73 using 27 out of 41 features.

4.2. Classification Results. The DNN model that was proposed achieved a greater level of accuracy, precision, recall, and F1 score than any of the previous papers while also requiring less computational time. According to the results presented in Table 4, our model had an accuracy of 99.73%, which was the greatest among the comparable research. Table 5 provides the confusion matrix that was created by applying the suggested model.

To examine the effectiveness of the proposed model, the results are compared with existing deep learning and

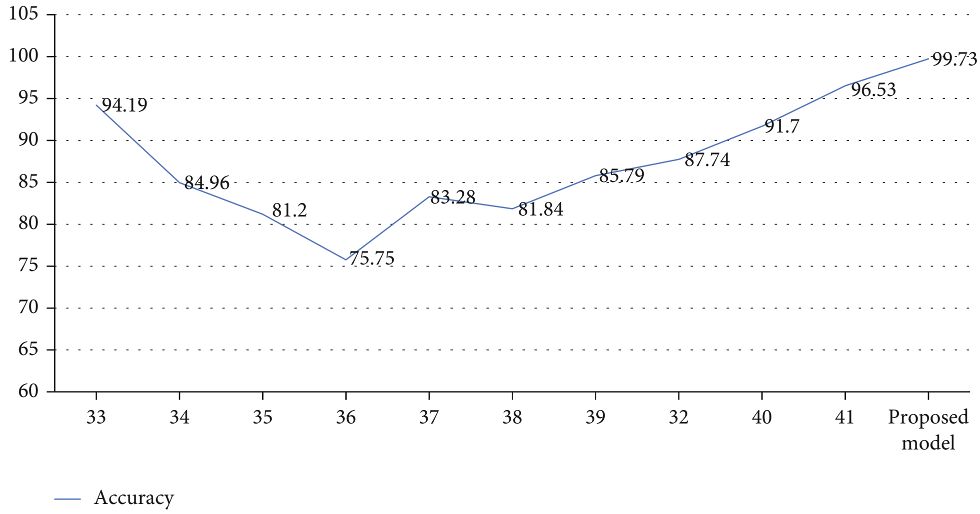

Figure 5: Accuracy comparison of the proposed model with other models.

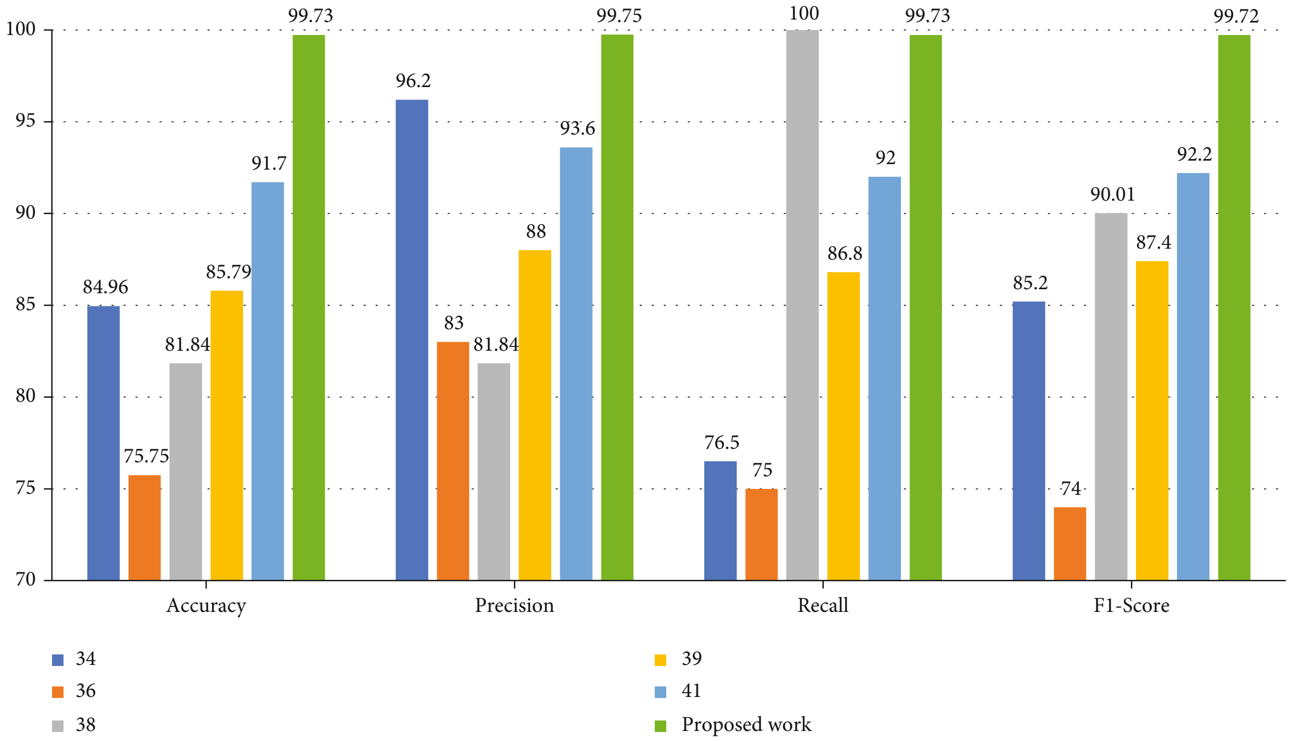

Figure 6: Performance of the proposed model and other models.

machine learning approaches. Table 6 demonstrates that the proposed model outperforms other benchmark algorithms.

Figures 5 and 6 show the evaluation results of the proposed model compared with the i^{th} exiting deep learning and machine learning models. Our proposed model achieves higher accuracy, precision, recall, and F1 score than all compared techniques. Our model also takes less training and testing than the comparative techniques.

Figure 7 shows the comparison of our proposed model with machine learning models. Our proposed model shows better results than all comparative machine learning techniques. Machine learning models are found to be struggling against network data.

4.2.1. Computational Time. In addition to other performance metrics, computational time is also an important metric that can be used to check the efficiency of the system for real situations of network intrusion detection. The Table 7 shows the training and testing time of our model compared with some other method time. Proposed method reduces the dimensionality of data; therefore, less computational power is required to train the model. The significance

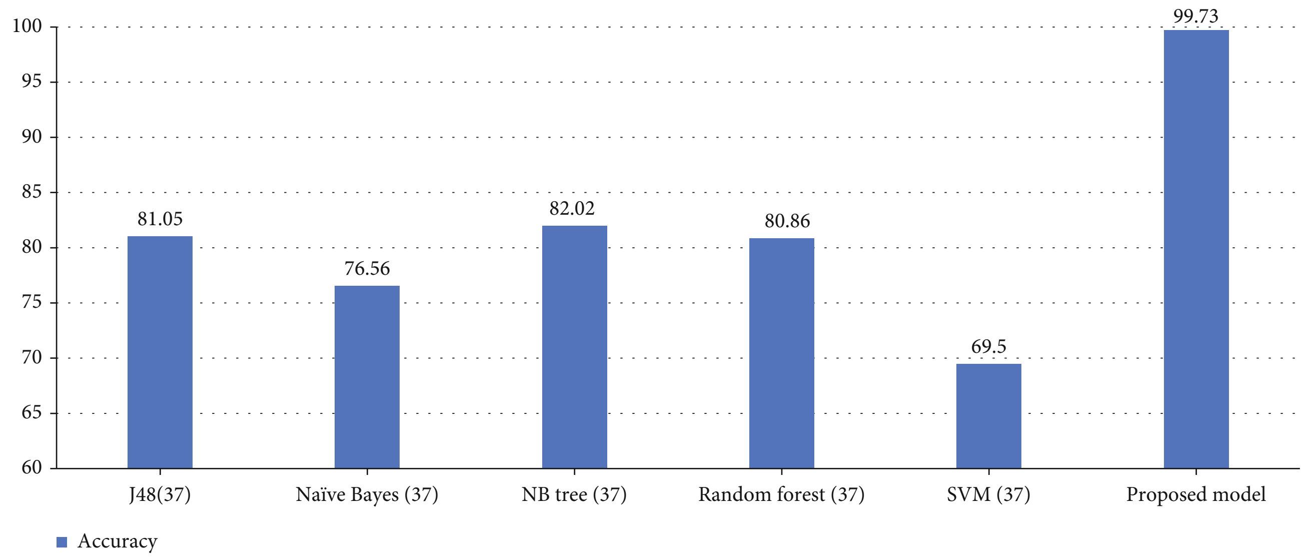

Figure 7: Comparison of the proposed work with machine learning models.

Table 7: Comparison computational time with other models.

Model	Training time (s)	Testing time (s)
Lakhina et al. [15]	40688	265
Potluri and Diedrich [26]	145.593437	—
Al-Qatf et al. [6]	673.031	4.648
Wang et al. [48]	1075s	—
Proposed model	138	2.7

of our model is the mitigation of overfitting effect due to the removal of redundant features using feature selection methods and reduction of time and computational complexity, whereas existing methods have only focused on the accuracy.

5. Conclusion and Future Work

The article presents a deep learning-based intrusion detection model. In the proposed scheme, the network data can be secured using the detection model and can save network data from all types of cyberattacks. The proposed model is the combination of feature selection classification techniques and achieves higher accuracy, precision, recall, and F1 score. The significance of our model is the mitigation of overfitting effect due to the removal of redundant features using feature selection methods and reduction of time and computational complexity, whereas existing methods have only focused on the accuracy. The proposed model also takes less time in training and testing than other comparative techniques. In future, this work can be applied to other datasets to check the performance. Similarly, we can also use a multiclass dataset to validate the model performance. For feature selection, we can use some other techniques which can improve the current model performance.

Data Availability

The data used in this research can be obtained from the corresponding authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R193), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

[1] S. Alabdulwahab and B. Moon, "Feature selection methods simultaneously improve the detection accuracy and model building time of machine learning classifiers," *Symmetry*, vol. 12, no. 9, p. 1424, 2020.

[2] S. H. Kok, A. Abdullah, N. Z. Jhanjhi, and M. Supramaniam, "A review of intrusion detection system using machine learning approach," *International Journal of Engineering Research and Technology*, vol. 12, no. 1, pp. 8–15, 2019.

[3] F. Shah, A. Anwar, H. AlSalman, S. Hussain, and S. Al-Hadhrami, "Artificial intelligence as a service for immoral content detection and eradication," *Scientific Programming*, vol. 2022, Article ID 6825228, 9 pages, 2022.

[4] C. Tang, N. Luktarhan, and Y. Zhao, "SAAE-DNN: deep learning method on intrusion detection," *Symmetry*, vol. 12, no. 10, p. 1695, 2020.

[5] X. Wang, S. Yin, H. Li, J. Wang, and L. Teng, "A network intrusion detection method based on deep multi-scale convolutional neural network," *International Journal of Wireless Information Networks*, vol. 27, no. 4, pp. 503–517, 2020.

[6] M. Al-Qatf, Y. Lasheng, M. Al-Habib, and K. Al-Sabahi, "Deep learning approach combining sparse autoencoder with SVM

for network intrusion detection," *IEEE Access*, vol. 6, pp. 52843–52856, 2018.

[7] B. Ingre and A. Yadav, "Performance analysis of NSL-KDD dataset using ANN," in *2015 international conference on signal processing and communication engineering systems*, pp. 92–96, Guntur, India, 2015, January.

[8] J. Kainat, S. Sajid Ullah, F. S. Alharithi, R. Alroobaea, S. Hussain, and S. Nazir, "Blended features classification of leaf-based cucumber disease using image processing techniques," *Complexity*, vol. 2021, Article ID 9736179, 12 pages, 2021.

[9] J. Iqbal, S. Hussain, H. AlSalman, M. A. Mosleh, and S. Sajid Ullah, "A computational intelligence approach for predicting medical insurance cost," *Mathematical Problems in Engineering*, vol. 2021, Article ID 1162553, 13 pages, 2021.

[10] Z. Li, Z. Qin, K. Huang, X. Yang, and S. Ye, "Intrusion detection using convolutional neural networks for representation learning," in *ICONIP 2017: Neural Information Processing, International Conference on Neural Information Processing*, Lecture Notes in Computer Science, Springer, Cham, 2017.

[11] B. A. Tama, M. Comuzzi, and K. H. Rhee, "TSE-IDS: a two-stage classifier ensemble for intelligent anomaly-based intrusion detection system," *IEEE Access*, vol. 7, pp. 94497–94507, 2019.

[12] S. Choudhary and N. Kesswani, "Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT," *Procedia Computer Science*, vol. 167, pp. 1561–1573, 2020.

[13] F. Farahnakian and J. Heikkonen, "A deep auto-encoder based approach for intrusion detection system," in *2018 20th international conference on Advanced communication technology (ICACT)*, pp. 178–183, Chuncheon, Korea (South), 2018, February.

[14] R. Boutaba, M. A. Salahuddin, N. Limam et al., "A comprehensive survey on machine learning for networking: evolution, applications and research opportunities," *Journal of Internet Services and Applications*, vol. 9, no. 1, pp. 1–99, 2018.

[15] S. Lakhina, S. Joseph, and B. Verma, *Feature Reduction Using Principal Component Analysis for Effective Anomaly–Based Intrusion Detection on NSL-KDD*, Citeseer, 2010.

[16] D. B. Rawat and S. R. Reddy, "Software defined networking architecture, security and energy efficiency: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 325–346.

[17] M. Z. Khan, R. Naseem, A. Anwar et al., "A novel approach to automate complex software modularization using a fact extraction system," *Journal of Mathematics*, vol. 2022, Article ID 8640596, 19 pages, 2022.

[18] I. Srba and M. Bieliková, "Encouragement of Collaborative Learning Based on Dynamic Groups," *21st Century Learning for 21st Century Skills*, A. Ravenscroft, S. Lindstaedt, C. D. Kloos, and D. Hernández-Leo, Eds., 2012.

[19] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, 2011.

[20] D. Hussain, M. Ismail, I. Hussain, R. Alroobaea, S. Hussain, and S. S. Ullah, "Face mask detection using deep convolutional neural network and MobileNetV2-Based transfer learning," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1536318, 10 pages, 2022.

[21] S. Kumar, A. Jain, S. Rani, H. Alshazly, S. A. Idris, and S. Bourouis, "Deep neural network based vehicle detection and classification of aerial images," *Intelligent Automation and Soft Computing.*, vol. 34, no. 1, pp. 119–131, 2022.

[22] M. Aa, M. Hamdi, S. Bourouis, K. Rastislav, and F. Mohmed, "Evaluation of neuro images for the diagnosis of Alzheimer's disease using deep learning neural network," *Frontiers in Public Health*, vol. 10, article 834032, 2022.

[23] A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues," *Knowledge-Based Systems*, vol. 189, p. 105124, 2020.

[24] K. Scarfone and P. Mell, "Guide to intrusion detection and prevention systems (idps)," *NIST Special Publication*, vol. 800, no. 2007, p. 94, 2007.

[25] N. Gao, L. Gao, Q. Gao, and H. Wang, "An intrusion detection model based on deep belief networks," in *2014 Second International Conference on Advanced Cloud and Big Data*, pp. 247–252, Huangshan, China, 2014, November.

[26] S. Potluri and C. Diedrich, "Accelerated deep neural networks for enhanced intrusion detection system," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp. 1–8, Berlin, Germany, 2016, September.

[27] J. Kim, N. Shin, S. Y. Jo, and S. H. Kim, "Method of intrusion detection using deep neural network," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 313–316, Jeju, 2017, February.

[28] Y. Zeng, H. Gu, W. Wei, and Y. Guo, "$ Deep-full-range $: a deep learning based network encrypted traffic classification and intrusion detection framework," *IEEE Access*, vol. 7, pp. 45182–45190, 2019.

[29] M. Maithem and G. A. Al-sultany, "Network intrusion detection system using deep neural networks," *In Journal of Physics: Conference Series*, vol. 1804, no. 1, p. 12138, 2021.

[30] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019.

[31] S. Jo, H. Sung, and B. H. Ahn, "A comparative study on the performance of SVM and an artificial neural network in intrusion detection," *Journal of the Korea Academia-Industrial Cooperation Society*, vol. 17, no. 2, pp. 703–711, 2016.

[32] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in *2016 international conference on wireless networks and mobile communications (WINCOM)*, pp. 258–263, Fez, Morocco, 2016, October.

[33] T. Su, H. Sun, J. Zhu, S. Wang, and Y. Li, "BAT: deep learning methods on network intrusion detection using NSL-KDD dataset," *IEEE Access*, vol. 8, pp. 29575–29585, 2020.

[34] Y. Yu and N. Bian, "An intrusion detection method using few-shot learning," *IEEE Access*, vol. 8, pp. 49730–49740, 2020.

[35] D. Kshirsagar and S. Kumar, "An efficient feature reduction method for the detection of DoS attack," *ICT Express*, vol. 7, no. 3, pp. 371–375, 2021.

[36] S. S. Ahmadi, S. Rashad, and H. Elgazzar, "Efficient feature selection for intrusion detection systems," in *2019 IEEE 10th annual ubiquitous computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 1029–1034, New York, NY, USA, 2019, October.

[37] X. Liu, T. Li, R. Zhang, D. Wu, Y. Liu, and Z. Yang, "A GAN and feature selection-based oversampling technique

for intrusion detection," *Security and Communication Networks*, vol. 2021, Article ID 9947059, 15 pages, 2021.

[38] O. Osanaiye, O. Ogundile, F. Aina, and A. Periola, "Feature selection for intrusion detection system in a cluster-based heterogeneous wireless sensor network," *Facta Universitatis, Series: Electronics and Energetics*, vol. 32, no. 2, pp. 315–330, 2019.

[39] F. Gottwalt, E. Chang, and T. Dillon, "CorrCorr: a feature selection method for multivariate correlation network anomaly detection techniques," *Computers & Security*, vol. 83, pp. 234–245, 2019.

[40] H. P. Vinutha and B. Poornima, "Analysis of feature selection algorithms for Naïve Bayes classifier using NSL-KDD," *International Journal of Engineering and Manufacturing Science*, vol. 8, no. 1, pp. 167–175, 2018.

[41] A. Dawoud, S. Shahristani, and R. Raun, "Deep learning and software defined networks: Towards secure IoT architecture," *Internet of Things*, vol. 3-4, pp. 82–89, 2018.

[42] S. Bhattacharya and S. Selvakumar, "LAWRA: a layered wrapper feature selection approach for network attack detection," *Security and Communication Networks*, vol. 8, no. 18, p. 3468, 2015.

[43] M. G. Raman, N. Somu, K. Kirthivasan, R. Liscano, and V. S. Sriram, "An efficient intrusion detection system based on hypergraph - genetic algorithm for parameter optimization and feature selection in support vector machine," *Knowledge-Based Systems*, vol. 134, pp. 1–12, 2017.

[44] S. Mohammadi, H. Mirvaziri, M. Ghazizadeh-Ahsaee, and H. Karimipour, "Cyber intrusion detection by combined feature selection algorithm," *Journal of Information Security and Applications*, vol. 44, pp. 80–88, 2019.

[45] X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, "An adaptive ensemble machine learning model for intrusion detection," *IEEE Access*, vol. 7, pp. 82512–82521, 2019.

[46] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.

[47] M. Torabi, N. I. Udzir, M. T. Abdullah, and R. Yaakob, "A review on feature selection and ensemble techniques for intrusion detection system," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, p. 2, 2021.

[48] Z. Wang, Y. Liu, D. He, and S. Chan, "Intrusion detection methods based on integrated deep learning model," *Computers & Security*, vol. 103, article 102177, 2021.

[49] S. M. Usman, S. Latif, and A. Beg, "Principle components analysis for seizures prediction using wavelet transform," *International Journal of Advanced and Applied Sciences*, vol. 6, no. 3, pp. 50–55, 2019.

[50] S. M. Usman, S. Khalid, R. Akhtar, Z. Bortolotto, Z. Bashir, and H. Qiu, "Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: review of available methodologies," *Seizure*, vol. 71, pp. 258–269, 2019.

[51] S. M. Usman, M. Usman, and S. Fong, "Epileptic seizures prediction using machine learning methods," *Computational and Mathematical Methods in Medicine*, vol. 2017, 10 pages, 2017.

[52] S. M. Usman, S. Khalid, and M. H. Aslam, "Epileptic seizures prediction using deep learning techniques," *IEEE Access*, vol. 8, pp. 39998–40007, 2020.

[53] S. M. Usman, S. Khalid, S. Jabbar, and S. Bashir, "Detection of preictal state in epileptic seizures using ensemble classifier," *Epilepsy Research*, vol. 178, article 106818, 2021.

[54] S. M. Usman, S. Khalid, and Z. Bashir, "Epileptic seizure prediction using scalp electroencephalogram signals," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 1, pp. 211–220, 2021.

[55] S. M. Usman, S. Khalid, and S. Bashir, "A deep learning based ensemble learning method for epileptic seizure prediction," *Computers in Biology and Medicine*, vol. 136, article 104710, 2021.

[56] A. Nazir and R. A. Khan, "A novel combinatorial optimization based feature selection method for network intrusion detection," *Computers & Security*, vol. 102, p. 102164, 2021.

[57] A. S. Dina and D. Manivannan, "Intrusion detection based on Machine Learning techniques in computer networks," *Internet of Things*, vol. 16, 100462 pages, 2021.

[58] F. Shah, Y. Liu, A. Anwar et al., "Machine learning: the backbone of intelligent trade credit-based systems," *Security and Communication Networks*, vol. 2022, Article ID 7149902, 10 pages, 2022.

[59] I. U. Haq, A. Anwar, I. U. Rehman et al., "Dynamic group formation with intelligent tutor collaborative learning: a novel approach for next generation collaboration," *IEEE Access*, vol. 9, pp. 143406–143422, 2021.

[60] C. Iwendi, P. K. Maddikunta, T. R. Gadekallu, K. Lakshmanna, A. K. Bashir, and M. J. Piran, "A metaheuristic optimization approach for energy efficiency in the IoT networks," *Software: Practice and Experience*, vol. 51, no. 12, pp. 2558–2571, 2021.

[61] C. Y. Chang, S. Bhattacharya, P. M. Raj Vincent, K. Lakshmanna, and K. Srinivasan, "An efficient classification of neonates cry using extreme gradient boosting- assisted grouped-support-vector network," *Journal of Healthcare Engineering*, vol. 2021, Article ID 7517313, 14 pages, 2021.

[62] R. Kaluri, D. S. Rajput, Q. Xin et al., "Roughsetsbased approach for predicting battery life in IoT," 2021, https://arxiv.org/abs/2102.06026.

[63] Z. Chiba, N. Abghour, K. Moussaid, A. El Omri, and M. Rida, "A novel architecture combined with optimal parameters for back propagation neural networks applied to anomaly network intrusion detection," *Computers & Security*, vol. 75, pp. 36–58, 2018.

[64] S. Ahmad, F. Arif, Z. Zabeehullah, and N. Iltaf, "Novel approach using deep learning for intrusion detection and classification of the network traffic," in *2020 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pp. 1–6, Tunis, Tunisia, 2020, June.