Predicting Cyber-attacks on IoT Networks using Deep-Learning and Different Variants of SMOTE

Bathini Sai Akash[1], Pavan Kumar Reddy Yannam[2], Bokkasam Venkata Sai Ruthvik[3], Lov Kumar[4], Lalita Bhanu Murthy[5], Prof. Aneesh Krishna[6] BITS-Pilani Hyderabad, India[1,2,3,4],Curtin University [5]
(,f20190065[1],f20190038[2],f20190017[3], lovkumar[4]
bhanu[5])@hyderabad.bits-pilani.ac.in,A.Krishna@curtin.edu.au[6]

Abstract. Predicting Cyber-attacks on IoT Networks using Machine Learning has a definite advantage over traditional methods because it helps secure against future attacks by identifying hidden patterns from past data, thereby improving the capability of a network. Thus automated systems are being developed which can be used to identify Cyber-attacks on IoT Networks using various machine learning techniques. In this work, three different types of features selection techniques were applied to the UNSW-NB15 data to find the best combination of relevant features. These selected sets of relevant features were considered to train five different deep learning architectures used to predict cyber attacks by varying the number of hidden layers. To handle the dataset's class imbalance problem, we have considered three different sampling techniques: SMOTE, Borderline SMOTE (BSMOTE), and Adaptive Synthetic Sampling (ADASYN). The experimental results on the UNSW-NB15 data highlight that the usage of considered feature selection techniques and class balance techniques does not significantly improve the predictive ability to detect cyber attacks. The results also suggest that variation in Deep learning Architecture impacts the prediction of cyberattacks.

Keywords: Cyber-attacks, Data Imbalance Methods, Feature Selection, Classification Techniques, Deep Learning.

1 Introduction

The use of IoT infrastructure has reached a new high. During the COVID-19 pandemic, there was a significant increase in the use of Smart Devices for domestic, medical, and industrial purposes[1]. The inevitable growth in the complexity of IoT infrastructures has raised unwanted vulnerabilities in various systems, which has led to a rise in Cyber-attacks on IoT Networks. This threat needs to be addressed since it could compromise the security of these networks and the infrastructure dependent on them[2][3]. For this purpose, Network intrusion detection systems (NIDS) are tasked with constantly surveying and detecting vulnerabilities simply by monitoring the network traffic. It was observed from the literature that Intrusion Detection Systems (IDS) use three different detection

methods: signature-based detection, anomaly-based detection, and a machine learning-based approach to identify a potential intrusion[4]. Machine Learning-based IDS helps to detect novel mutated attacks by learning from heterogeneous data, and earlier forms of attack [5]. Consequently, researchers working in the area of IDS observed that Machine Learning based IDS is more resilient and does not need as many patches as in traditional techniques like signature-based detection [6].

Hence, using Machine Learning has a definite advantage because it helps stop future attacks on IoT Networks and thereby improves the capability of a network to handle future attacks by identifying the pattern of attacks. In order to detect cyber-attacks on IoT networks, several machine learning methods were considered in the literature, but low importance was given to finding the optimal combination of feature selection techniques and class balancing techniques to be used along with Deep Learning methods. Deep Learning techniques can identify hidden patterns in large amounts of data, i.e., big data. This study made use of the UNSW-NB15 [7][8] dataset for network intrusion detection to validate the proposed framework, with over 170,000 training records. The dataset contains modern network traffic, including low footprint attack types. However, it was observed that the considered dataset suffers from the class imbalance problem. A dataset is defined as a balanced dataset when roughly an equal number of input samples represents each target class. So, this study used three different data sampling techniques to balance the data with an objective to improve the predictive ability of the developed models. Finally, we also used different feature selection techniques to remove unimportant features. This helps reduce the computational complexity of the models and potentially improve the performance of Machine learning algorithms by extracting only the necessary features[9]. The key contributions of this research are as follows:

- Five different types of deep learning architectures were considered to develop a model for predicting Cyber-attacks on IoT Networks.
- Three different data sampling techniques were used to balance the data before applying the training algorithm.
- Three different feature selection methods were also applied to select the right combination of features with an objective to better the performance of deep learning models.

2 Related Work

K.A Tait et al. [10] provided a comparison of various machine learning techniques applied on the UNSW-NB15 data set. They reported that the random forest algorithm performed better than other methods in the case of binary classification, achieving an accuracy of 99.77%, while KNN gave the best results in the case of multi-class classification. C.Wheelus et al. [11] proposed the prepossessing step to tackle the class imbalance problem in cyber security datasets such as UNSW NB15 [7][8] and KDD-CUP99 [12]. The study re-processed the UNSW

NB15 dataset using the attributes from the SANTA data set. They reported that SMOTE provided the most significant lift on average in the ROC-AUC values, compared to the baseline values. Nutan Farah Haq et al. [13] have analyzed 49 research papers related to using different intrusion detection classifiers between 2009-2014 and observed the need to remove redundant and irrelevant features for the training phase while stating that it is a crucial factor for system performance. Fatemeh Amiri et al. [14] exclusively considered the effect of feature selection techniques such as the forward feature selection algorithm, linear correlation-based feature selection, and a modified mutual information feature selection algorithm in preprocessing the KDD-CUP99 dataset. Finally, they have concluded that the feature selection algorithms can significantly improve classification accuracy.

To the best of our knowledge, all the related works in this field have notably used deep learning for cyber-attacks on IoT Networks. Our work analyses deep learning methods coupled with different feature selection and sampling techniques to predict cyber-attacks on IoT Networks.

3 STUDY DESIGN

This section presents the details regarding the various design settings used for this research.

3.1 Experimental Dataset

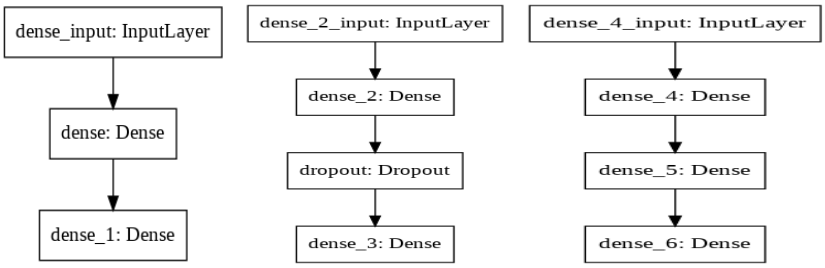
This study used the UNSW-NB15 dataset, published in 2015 [7][8]. It has 49 features comprising nine modern attacks and class labels, synthetically produced in a test environment. Attack types such as Worms, Shellcode, Generic, Reconnaissance, DoS exploits, Fuzzers, and Backdoors are provided for multi-class classification. This study chose the UNSW-NB15 dataset over other recognized datasets such as the KDD CUP99 [12] or the NSL-KDD because it contains modern network traffic of normal and abnormal types, including low footprint attack types. In contrast, the other datasets mentioned above are relatively older and may not represent modern network traffic.

3.2 Feature Selection Techniques

Feature selection methods facilitate a way of reducing computation time and avoid the curse of dimensionality. This study analyzes whether these techniques improve prediction performance or mask critical features in the network intrusion detection domain. This work made use of three different feature selection techniques: Significance test, Cross-Correlation test, and PCA[15] with an objective to remove irrelevant features and select the most appropriate set of features for predicting cyber-attacks on IoT Networks.

3.3 Deep Learning Architecture

Deep Learning (DL) is widely known for finding complex patterns in data[1]. Deep Learning makes use of non-linear processing units, called neurons, stacked to form multiple layers, which collectively combine the input signals in complex ways to get an appropriate output.[16] This paper used five different types of deep learning architectures by changing the number of hidden layers with an objective to develop models for detecting cyber-attacks on IoT Networks. The architecture of the considered DL models (DL1, DL2, DL3, DL4, and DL5) is shown in figure 1. The considered cyber-attack prediction models were trained using a batch size of 128 and an early stopper which allowed a maximum of 1000 epochs. The dropout hyperparameter value used was 0.2.



(1.1) DL1 (1.2) DL2 (1.3) DL3

Fig. 1: Deep Learning Architectures

3.4 SMOTE

The above-proposed models were validated with the help of the UNSW-NB15 dataset as mentioned in section 3. However, the UNSW-NB15 dataset does not have an equal representation of attack and benign samples; in this case, only 1 out of 3 data points have a positive label and hence are classified under the attack category. So, this study used three sampling techniques: SMOTE, Borderline SMOTE (BSMOTE/BLSMOTE), and Adaptive Synthetic Sampling (ADASYN)[17] to overcome the problem of data imbalance[11]. In this work, we also validated the predictive ability of the models trained on the balanced data and original data(OD).

4 Results and analysis

This section presents the performance of the trained models generated using different feature selection techniques, data sampling techniques, and different architectures of deep-learning models in terms of Accuracy and AUC scores. Figure 2 shows the steps followed to develop the prediction models for cyber-attacks on IoT Networks. These steps are as follow:

– Countvectorizer was applied to the dataset to convert the text features of the dataset to numerical features.
– The preprocessed data after the above step was passed as an input for different feature selection techniques to find the best set of features for predicting cyber-attacks on IoT Networks.
– Different data sampling techniques have been applied on the selected features with an objective to get an equal distribution of samples belonging to both classes.
– The resulting final preprocessed data was used as an input to five different deep-learning models with an objective to find patterns that help identify future cyber-attacks on IoT Networks.
– The models developed using the above techniques were validated with the help of the UNSW-NB15 data set.

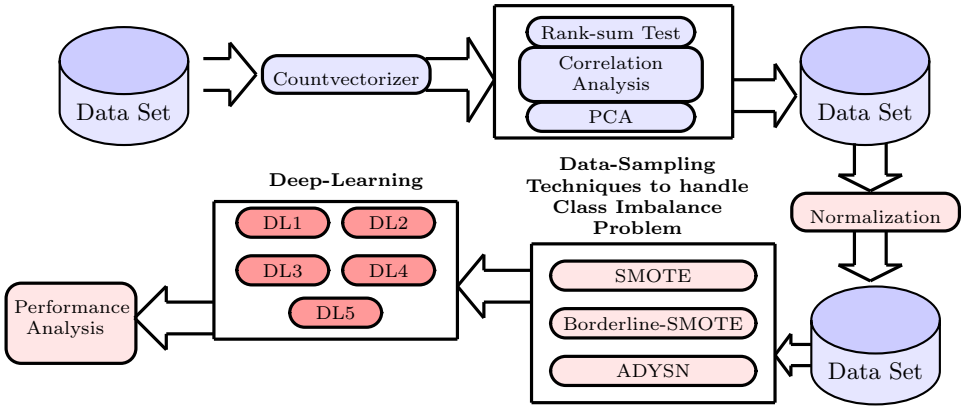


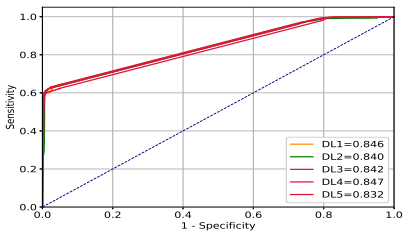
Fig. 2: Proposed Framework

Three different combinations of features, four different sets of data, and five different variants of deep learning have been used to develop models for predicting cyber-attacks on IoT Networks. Therefore, a total of 120 ((3 sets of features)* (1PCA+1 Without PCA)*(1 Original Data+ 3 class balancing techniques)* 5 different classification techniques) distinct prediction models were generated in the study. The performance of these models is evaluated in terms of Accuracy, F-Measure, and AUC scores. Figure 3 and Table 1 show the performance of models trained on preprocessed data that involved PCA. The results for other cases are of similar type. The information present in Table 1 and Figure 3 suggested the value of AUC$\geq$0.7 for all models confirms that the developed models have the ability to predict future cyber-attacks on IoT Networks. The information also suggested that the models developed using DL3 have a better ability to predict cyber-attacks on IoT Networks than others.

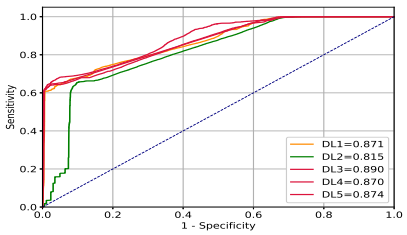
5 Comparative Analysis

This section presents the performances of various models obtained for intrusion detection. As explained in section 3, this paper applies various techniques such

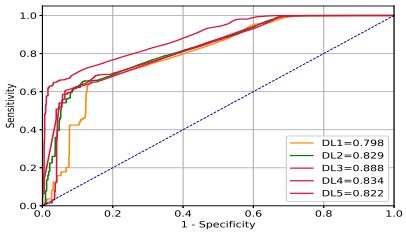
as class balancing and feature selection on the UNSW-NB15 dataset to validate the proposed models. We have also investigated whether neural networks with different hidden layers and dropout regularization have any significant difference in their performance for the classification problem of intrusion detection. The results obtained in this study are summarized in the following subsections:



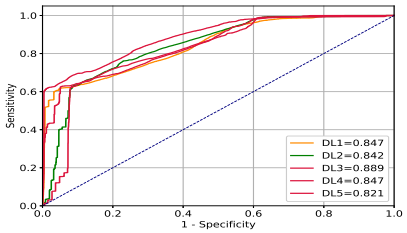
(3.1) SIGF+PCA (BSMOTE)



(3.2) AF+PCA (ADSYN)



(3.3) SIGF+PCA (ADSYN)



(3.4) CCRA+PCA (ADSYN)

Fig. 3: ROC Curve

Table 1: Accuracy, F-Measure, and AUC Values

		Accuracy					**F-Measure**					**AUC**				
		DL1	**DL2**	**DL3**	**DL4**	**DL5**	**DL1**	**DL2**	**DL3**	**DL4**	**DL5**	**DL1**	**DL2**	**DL3**	**DL4**	**DL5**
OD	AF+PCA	70.00	71.16	73.54	72.63	72.78	0.78	0.79	0.79	0.79	0.78	0.82	0.82	0.87	0.87	0.83
OD	SIFG+PCA	70.29	69.20	70.73	71.95	69.50	0.78	0.78	0.78	0.79	0.78	0.83	0.84	0.84	0.83	0.82
OD	CCRA+PCA	71.23	67.35	71.92	71.71	71.02	0.74	0.75	0.79	0.79	0.78	0.80	0.80	0.84	0.84	0.83
SMOTE	AF+PCA	74.27	75.11	75.55	71.38	76.41	0.77	0.75	0.77	0.78	0.79	0.87	0.83	0.87	0.86	0.88
SMOTE	SIFG+PCA	73.08	74.77	73.35	71.96	71.61	0.75	0.75	0.76	0.78	0.77	0.83	0.83	0.85	0.84	0.83
SMOTE	CCRA+PCA	72.04	73.51	72.46	74.02	77.52	0.74	0.75	0.78	0.74	0.75	0.83	0.85	0.85	0.80	0.86
BSMOTE	AF+PCA	78.48	75.58	79.96	79.26	78.71	0.77	0.75	0.78	0.77	0.76	0.88	0.83	0.89	0.87	0.86
BSMOTE	SIFG+PCA	74.43	74.19	76.47	75.18	73.79	0.74	0.74	0.75	0.74	0.73	0.83	0.83	0.86	0.84	0.83
BSMOTE	CCRA+PCA	74.99	74.80	77.53	73.16	78.02	0.75	0.74	0.76	0.74	0.75	0.82	0.84	0.88	0.83	0.76
ADSYN	AF+PCA	78.24	74.79	79.82	78.93	78.81	0.77	0.74	0.78	0.77	0.76	0.87	0.82	0.89	0.87	0.87
ADSYN	SIFG+PCA	74.09	73.86	78.87	73.86	73.90	0.74	0.74	0.78	0.75	0.73	0.80	0.83	0.89	0.83	0.82
ADSYN	CCRA+PCA	75.07	75.57	77.90	75.55	75.46	0.74	0.75	0.76	0.75	0.74	0.85	0.84	0.89	0.85	0.82

5.1 Sampling Techniques

We have used three different types of sampling techniques: SMOTE, Borderline SMOTE, and ADASYN aiming to make an equal distribution of samples in both the categories. In this section, we aim to find the solution for ***RQ1: "Is there***

a significant difference between models generated using various class balancing techniques?" based on box plots and significant tests.

Box-Plot: Sampling Techniques: The performances of the models developed using different types of sampling techniques in terms of accuracy, F-measure, and AUC are represented in Figure 4. The information present in Figure 4 suggested that the models developed using original data have a better ability to predict cyber-attacks on IoT networks as compared to sampling techniques. As mentioned in Figure 4, the models developed using original data achieved an average AUC value of 0.91, while the models developed using balanced data achieved an average AUC value of 0.82.

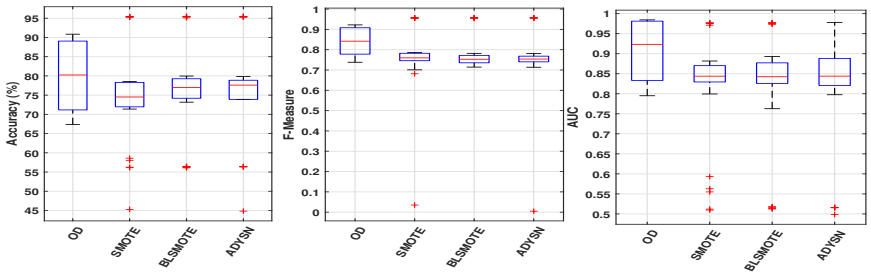


Fig. 4: Box-Plot Diagram: Performance of Different Sampling Techniques

Statistical Hypothesis Testing: In this work, we used statistical hypothesis testing to validate our considered null hypothesis, i.e., "There is no significant difference in the performance of the models across all the different class balancing methods." The above null hypothesis is accepted if the calculated p-value is greater than 0.05. The results of statistical hypothesis testing on the performance of different models trained using different data balancing techniques are mentioned in Table 2. From Table 2, it can be observed that all comparisons between SMOTE, BSMOTE, and ADASYN have P-values greater than 0.05, which means that the null hypothesis cannot be rejected in any pairing. This indicates that there is no significant difference in the performance of the models across all the different class balancing methods applied.

Table 2: Hypothesis Testing: Different Sampling Techniques

	OD	**SMOTE**	**BSMOTE**	**ADASYN**
OD	1.000	0.022	0.018	0.021
SMOTE	0.022	1.000	0.994	0.923
BSMOTE	0.018	0.994	1.000	0.947
ADASYN	0.021	0.923	0.947	1.000

5.2 Feature Selection Techniques

In this work, we have used three different feature selection techniques: Significance test-based feature selection (SIGF), features selected after cross-correlation analysis, and PCA with an objective to remove irrelevant features and select the

right combination of features. In this section, we are going to find the solution for ***RQ2: "What is the impact of using feature selection techniques for developing effective models?"*** based on box-plot and significant test.
Box-Plot: Sampling Techniques: The performance of the models developed using different types of feature selection techniques in terms of Accuracy, F-measure, and AUC are visualized as a box plot in Figure 5. The information present in Figure 5 suggests that the models developed using Actual Features (AF - Baseline) have better ability to predict cyber-attacks on IoT networks as compared to feature selection techniques. The models developed using original data achieved an average AUC value of 0.98, while the models developed using feature selection techniques achieved an average AUC value of 0.76.

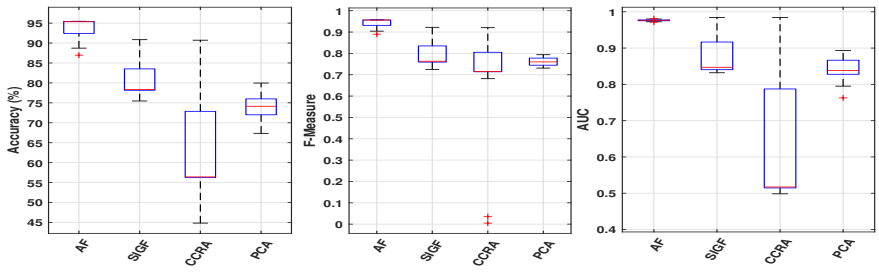


Fig. 5: Box-Plot Diagram: Performance of Different Features Selection Techniques

Statistical Hypothesis Testing: Different Features Selection Techniques: In this work, we used statistical hypothesis testing to validate our considered null hypothesis, i.e., "The AUC performance value of the models developed for IDS prediction using neural networks used on datasets generated using different feature selection techniques is not significantly different." The above null hypothesis is accepted if the calculated p-value is greater than 0.05. The results of statistical hypothesis testing on the performance of different models trained using different feature selection techniques are mentioned in Table 3. From Table 3, it can be observed that the p-values are less than 0.05 for the pairings where one of SIGF, CCRA, and PCA are involved. This indicates that model trained using different features are significantly different.

Table 3: Hypothesis Testing: Different Features Selection Techniques

	AF	**SIGF**	**CCRA**	**PCA**
AF	1.00	0.00	0.01	0.00
SIGF	0.00	1.00	0.00	0.02
CCRA	0.01	0.00	1.00	0.00
PCA	0.00	0.02	0.00	1.00

5.3 Deep Learning Technique

We have used five different types of deep learning architectures by changing the number of hidden layers with an objective to develop models for detecting cyber-attacks on IoT Networks. In this section, we are going to find the solution for ***RQ2: "What is the impact of deep learning architecture on the performance of the models?"*** based on box plots and significant tests.
Box-Plot: Deep Learning Technique: The performance of five different types of deep learning models in terms of Accuracy, F-measure, and AUC are visualized as box plots in Figure 6. The information present in Figure 6 suggest that all the deep learning model Architectures have a similar ability to predict cyber-attacks on IoT networks. The average AUC value of DL3 - (Two Hidden Layers) is slightly higher than other proposed architectures. Therefore, based on descriptive statistics DL3 Model performs better than other proposed models for cyber attack prediction.

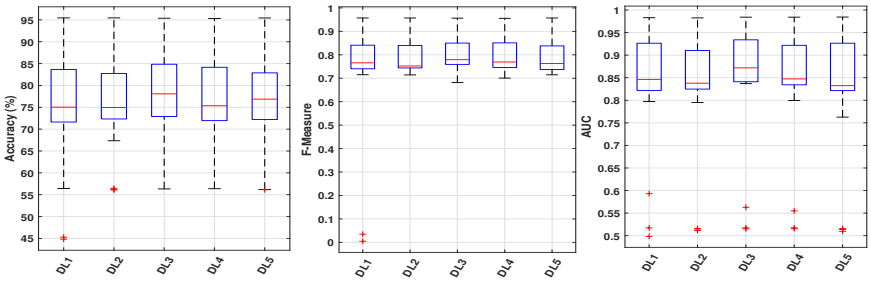


Fig. 6: Box-Plot Diagram: Performance of Different Deep-Learning Models

Statistical Hypothesis Testing: Different Deep-Learning Models: In this work, we have used statistical hypothesis testing to validate our considered null hypothesis, i.e., "The AUC performance value of the models developed for IDS prediction using neural networks used on datasets generated using DL2 and DL3 are significantly different."The above null hypothesis is accepted if the calculated p-value is greater than 0.05. The results of statistical hypothesis testing on the performance of different models trained using different feature selection techniques are mentioned in Table 4. From Table 4, it can be observed that the p-values are less than 0.05 for the pairing with DL2 and DL3. This, coupled with the descriptive statistics values, which indicate that the AUC value of DL3 is the highest, implies that DL3 has the best performance. DL3 is a model of two hidden layers, while DL2 has only one hidden layer.

5.4 Cost benefit analysis

This section presents the experimental setup used to find the effectiveness of the developed models using a cost-benefit analysis. The information presented in Table 1 suggested that the models trained using all features as an input have a better ability to predict attack as compared to the features selected using

Table 4: Hypothesis Testing: Different Deep-Learning Models

	DL1	**DL2**	**DL3**	**DL4**	**DL5**
DL1	1.00	0.73	0.15	0.53	0.84
DL2	0.73	1.00	0.04	0.22	0.93
DL3	0.15	0.04	1.00	0.23	0.10
DL4	0.53	0.22	0.23	1.00	0.30
DL5	0.84	0.93	0.10	0.30	1.00

feature selection techniques. Also, computation time increases with an increase in the number of features processed as time complexity increases. Hence, the cost-benefit analysis has been used to find the best combination of features in this work. The formula used in this paper is given below. Further, the amount of space the data occupies is also directly proportional to the number of features present. Since time and space are crucial resources while training a model, they constitute the cost of training a model. The benefit of a model is how effectively it can categorize attacks and differentiate categories into their respective classes. This directly translates to a more efficient intrusion detection system(IDS). The formula used to calculate the cost and benefit of the developed models is mentioned below:

$$Cost_{M1} = \lambda_f * \frac{AF - SF}{AF} + \lambda_a * AUC_{M1} \tag{1}$$

Where λ_f and λ_a are the weightage given to the number of features and performance of the models, respectively. Similarly, AF and SF are used to represent the number of all features and selected sets of features. In this work, we calculated the cost of models by varying the value of λ_f and λ_a. The result of cost-benefit analysis for different combinations of λ_f and λ_a are summarized in the table 5. For example, when we assign 50% importance to the reduction in cost (λ_f=0.5) and 50% (λ_a=0.5) importance to the AUC value(the benefit), the models trained on feature exacted after applying PCA on AF have high benefit as compared to others.

6 Conclusion

Modern IDS are capacitated in providing a wide range of services. However, even though significant work has been done in the field of IDS, they are still immature in many aspects, and certain problems are yet to be resolved. As a result, this research was on a Machine Learning-based IDS that can detect novel mutation attacks by learning from diverse data and previous forms of attack to make accurate predictions. Hence, this research study focused on evaluating whether specific methods in the preprocessing framework and Deep Neural Networks' architecture have a statistically significant impact on the performance of IDS compared to others. An empirical analysis was done using Descriptive Statistics and hypothesis testing to compare the performances of models generated across

Table 5: Cost-Benefit Analysis

λ_f	Rank 1	Rank 2	Rank 3
0.05	CCRA	SIGF	AF
0.1	CCRA	SIGF	AF
0.15	CCRA	SIGF	AF+PCA
0.2	CCRA	SIGF	AF+PCA
0.25	CCRA	SIGF	AF+PCA
0.3	CCRA	SIGF	AF+PCA
0.35	CCRA	AF+PCA	AF+PCA
0.4	CCRA	AF+PCA	AF+PCA
0.45	CCRA	AF+PCA	AF+PCA
0.5	AF+PCA	AF+PCA	CCRA
0.55	AF+PCA	AF+PCA	CCRA+PCA
0.6	AF+PCA	CCRA+PCA	AF+PCA
0.65	CCRA+PCA	AF+PCA	AF+PCA
0.7	CCRA+PCA	AF+PCA	AF+PCA
0.75	CCRA+PCA	CCRA+PCA	SGF+PCA
0.8	CCRA+PCA	CCRA+PCA	SGF+PCA
0.85	CCRA+PCA	CCRA+PCA	SGF+PCA
0.9	CCRA+PCA	CCRA+PCA	SGF+PCA
0.95	CCRA+PCA	CCRA+PCA	SGF+PCA

three data sampling techniques, three feature selection techniques, and different deep learning architectures. Our observations are the following:

- Our experiment results suggest that an increase in the number of hidden layers in Deep Neural Networks improves model performance.
- Our experiment results suggest that class balancing techniques such as SMOTE, Borderline SMOTE, and ADASYN do not significantly improve model performance.
- Our experiment results also suggest that significance test-based feature selection, cross correlation-based feature selection techniques, or PCA negatively affect model performance compared to the baseline of no feature selection (AF).
- The results of the cost-benefit analysis suggested that the models developed using PCA are computationally more efficient while giving reasonable performance compared to the case when all features are used.

7 Acknowledgment

This research is funded by TestAIng Solutions Pvt. Ltd.

References

1. Krishna Kumar, Narendra Kumar, and Rachna Shah. Role of iot to avoid spreading of covid-19. *International Journal of Intelligent Networks*, 1:32–35, 2020.

2. Mahdis Saharkhizan, Amin Azmoodeh, Ali Dehghantanha, Kim-Kwang Raymond Choo, and Reza M Parizi. An ensemble of deep recurrent neural networks for detecting iot cyber attacks using network traffic. *IEEE Internet of Things Journal*, 7(9):8852–8859, 2020.
3. Ioannis Stellios, Panayiotis Kotzanikolaou, Mihalis Psarakis, Cristina Alcaraz, and Javier Lopez. A survey of iot-enabled cyberattacks: Assessing attack paths to critical infrastructures and services. *IEEE Communications Surveys & Tutorials*, 20(4):3453–3495, 2018.
4. VVRPV Jyothsna, Rama Prasad, and K Munivara Prasad. A review of anomaly based intrusion detection systems. *International Journal of Computer Applications*, 28(7):26–35, 2011.
5. Mohammed Ali Al-Garadi, Amr Mohamed, Abdulla Khalid Al-Ali, Xiaojiang Du, Ihsan Ali, and Mohsen Guizani. A survey of machine and deep learning methods for internet of things (iot) security. *IEEE Communications Surveys & Tutorials*, 22(3):1646–1685, 2020.
6. Amiya Kumar Sahu, Suraj Sharma, M Tanveer, and Rohit Raja. Internet of things attack detection using hybrid deep learning model. *Computer Communications*, 2021.
7. Nour Moustafa and Jill Slay. Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). In *2015 military communications and information systems conference (MilCIS)*, pages 1–6. IEEE, 2015.
8. Nour Moustafa and Jill Slay. The evaluation of network anomaly detection systems: Statistical analysis of the unsw-nb15 data set and the comparison with the kdd99 data set. pages 1–14, 01 2016.
9. Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1), 2014.
10. Kathryn-Ann Tait, Jan Sher Khan, Fehaid Alqahtani, Awais Aziz Shah, Fadia Ali Khan, Mujeeb Ur Rehman, Wadii Boulila, and Jawad Ahmad. Intrusion detection using machine learning techniques: An experimental comparison. *arXiv preprint arXiv:2105.13435*, 2021.
11. Charles Wheelus, Elias Bou-Harb, and Xingquan Zhu. Tackling class imbalance in cyber security datasets. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pages 229–232. IEEE, 2018.
12. Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A Ghorbani. A detailed analysis of the kdd cup 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications*, pages 1–6. IEEE, 2009.
13. Nutan Farah Haq, Abdur Rahman Onik, Md Avishek Khan Hridoy, Musharrat Rafni, Faisal Muhammad Shah, and Dewan Md Farid. Application of machine learning approaches in intrusion detection system: a survey. *IJARAI-International Journal of Advanced Research in Artificial Intelligence*, 4(3):9–18, 2015.
14. Fatemeh Amiri, MohammadMahdi Rezaei Yousefi, Caro Lucas, Azadeh Shakery, and Nasser Yazdani. Mutual information-based feature selection for intrusion detection systems. *Journal of Network and Computer Applications*, 34(4):1184–1199, 2011.
15. XIAOHUI XIE. Principal component analysis. 2019.
16. Chuanlong Yin, Yuefei Zhu, Jinlong Fei, and Xinzheng He. A deep learning approach for intrusion detection using recurrent neural networks. *Ieee Access*, 5:21954–21961, 2017.
17. Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.