International Journal of Women in Technical Education and Employment
ISSN: 2811-1567. Volume 4 – Issue 2. December 2023
https://fpiwitedjournal.federalpolyilaro.edu.ng

INTERNATIONAL JOURNAL OF WOMEN IN TECHNICAL EDUCATION AND EMPLOYMENT (IJOWITED)

Performance Evaluation of Homogenous Boosting Technique for Online Banking Network Intrusion Detection

[1]Folahan Jiboku, [2]Wumi Ajayi & [3]Kayode Oladapo
[1]Department of Computer science, The Federal Polytechnic Ilaro, Nigeria, [2]Department of Computer Science, Babcock university, Ilishan-Remo, Ogun state, Nigeria, [3]Department of Physical and Computer Sciences, McPherson University, Seriki Sotayo, Ogun State, Nigeria. folahan.jiboku@federalpolyilaro.edu.ng; ajayiw@babcock.edu.ng;oladapoka@mcu.edu.ng

ARTICLE HISTORY

Received: Nov 24, 2023
Revised: Nov 30, 2023
Accepted: Dec 7, 2023

Abstract

In recent times, it has been observed that a lot of users have been using online banking. However, security of online banking has been a matter of great concern for most users. This paper presents a performance evaluation of a homogeneous boosting technique for online banking network intrusion detection. The study aims to determine the effectiveness of the boosting technique in improving the detection of network intrusion attempts in online banking systems. The research methodology includes applying fuzzy logic feature selection technique on the dataset to determine the objectivity of the homogenous boosting ensemble machine learning algorithms. The experimental results of the study showed that the homogenous boosting technique performed well on the datasets, achieving high levels of accuracy and recall. The study also showed that the homogeneous boosting technique has a relatively low false-positive rate, indicating a high level of precision in detecting network intrusion attempts. Furthermore, the study evaluated the impact of various feature selection techniques on the performance of the boosting technique. The results demonstrate that the boosting technique performed better with selected feature subsets, which implies that the technique can be optimized for different online banking network intrusion detection scenarios. In conclusion, this paper demonstrated the effectiveness of the homogeneous boosting technique for online banking network intrusion detection. The study provides valuable insights into the use of boosting techniques and feature selection for improving the detection of network intrusion attempts in online banking systems. The findings of this study could help enhance the security of online banking systems and improve the overall trust of customers in online banking.

Keywords: Online Banking, Intrusion Detection, Fuzzy Logic, Homogenous boosting.

Citation
Ajayi, W, Jiboku, F. & Oladapo, K. (2023). Performance Evaluation of Homogenous Boosting Technique for Online Banking Network Intrusion Detection. *International Journal of Women in Technical Education and Employment*, 4(2), 62 – 75.

Introduction

Online banking, often known as internet banking, e-banking, or virtual banking, is an electronic payment system that allows clients of a bank or other financial institution to execute a variety of financial transactions via the bank's website. The concept of online banking emerged in the late 20th century with the proliferation of the internet. In the mid-1990s, financial institutions started offering basic online services such as checking balances and viewing transaction history through web portals. In most cases, an internet banking system will link to or be a component of a bank's core banking system on a network (Abualsaoud et al., 2020).

To many people, electronic banking entails direct deposit of paychecks into checking or savings accounts or 24-hour access to cash through an automated teller machine (ATM). Online banking has

revolutionized the way people conduct financial transactions. It allows customers to access their accounts, make payments, transfer funds, and perform various other banking activities through the internet. This convenience, however, comes with its own set of challenges, primarily in the form of security threats. Intrusion detection plays a critical role in safeguarding online banking systems against these threats. The evolution of online banking and the pivotal role of intrusion detection systems in ensuring its security.

Over time, online banking evolved to include a wide range of services, including electronic funds transfers, bill payments, and investment management. This expansion was driven by advancements in internet technologies, encryption protocols, and the development of secure communication channels. However, there are numerous different sorts of transactions, rights, obligations, and occasionally fees, associated with using electronic banking. A variety of banking and other services or facilities that employ electronic technology are referred to as "banking services. "These consist of telephone banking, SMS banking, ATM and debit card services, electronic alerts, mobile banking, money-transfer services, point-of-sale banking, e-statements, and other forms of e-commerce or services that create value (Monil et al., 2020).

Online banking has a lot of advantages. The two most crucial ones are convenience and speed. Online banking users get access to their accounts, statements, transactions, bill payment options, and more from the comfort of their homes or while on the road. These advantages are the main reasons why about 51% of EU adults utilize internet banking. Online banking does have its advantages, but there are also several unique problems and difficulties in the industry. These are extremely important for banks that provide online banking as well as for their clients, who depend on the banks' smooth operation (Hammoud et al., 2018).

While electronic banking benefits the financial system, it poses significant security risks to institutions and their clients. Before a user may access bank services, they must first enter an access code, which is usually in the form of a Personal Identification Number (PIN). This has not always protected banks from fraudsters' theatrics; fraudsters utilize a variety of methods to reveal or steal clients' secret access numbers (Charkhar &Jagdeesh, 2018). Banks typically use manual inspection along with rules-based fraud detection technologies to find scams.

A security system should be able to guard itself against external intrusions; otherwise, fraudsters may choose to attack the system by turning it off (Lin et al., 2012). As a result, it's crucial that an electronic banking application has some level of security intelligence and can protect itself against existential threats or intrusions. The Intrusion Detection System (IDS) is a powerful protection tool against both network and host-based threats. It gathers, analyzes, and audits security logs and network packets while monitoring important nodes of computer systems or networks. An IDS focuses on the proactive and timely detection of external attackers and unusual server behavior before they do such severe damage. Several cyberattacks have been in dangerous situations as of late, putting certain organizations' vital infrastructures at risk. A successful attack may have unfavorable effects, including but not limited to financial loss, the end of operations, and the revealing of secret information. Additionally, attackers have a greater possibility of success the larger the organization's network is. The network's complexity may also result in weaknesses and other specialized threats. As a result, security mitigation and protection techniques ought to be viewed as required (Khan, 2021).

A potential defense system, like intrusion detection, is essential since it uses preventive measures to get rid of any malicious activities within the computer network. An IDS looks at network and file access logs, audit trails, and other security-relevant data within the organization to detect and block threats without human interaction (FBI, 2021).

Banks are now able to identify transactions that are most likely to be fraudulent while maintaining acceptable levels of false positives thanks to machine learning. The ensemble technique combines the results of various classifiers to create a single response (Onu et al., 2017). This strategy helps to achieve greater detection than the classification accuracy of a

single classifier. Several trainable classifiers, such as base learners, form the foundation of an ensemble learner. Each base learner has been taught to predict for a specific class label, with the final prediction being formed using a specific blending method, such as a combiner. It is assumed that classifier ensembles now outperform individual classifiers for a variety of reasons, including statistical, computational, and representational ones (Khan, 2021). Most studies on combining classifiers within the purview of IDS were initially started with a single justification.

The objective of this study is to design a fuzzified classifier model based on homogenous boosting ensemble machine learning algorithm for improved detection of intrusion through performance evaluation. Firstly, apply a fuzzy logic feature selection technique on the selected dataset to determine the objectivity of the homogenous boosting ensemble machine learning algorithms for the performance evaluation of intrusion detection model. Secondly, evaluate the performance of the homogenous boosting machine learning algorithms using the selected metrics, and lastly perform a comparative analysis of homogeneous boosting ensemble algorithm based on the evaluation criteria.

The proposed study will evaluate the effectiveness of the Homogeneous Boosting technique for online banking network intrusion detection. This technique combines multiple weak classifiers to create a strong classifier capable of accurately detecting intrusion attempts in real-time. The study's findings will help to determine whether the Homogeneous Boosting technique is a viable solution for improving the accuracy and efficiency of online banking network intrusion detection. Given the escalating sophistication of cyber threats targeting online banking systems, evaluating the efficacy of the homogenous boosting technique in intrusion detection addresses a critical need. The study's findings can lead to enhanced security protocols, thereby fortifying the resilience of online banking networks against evolving threats.

Methodology

The methodology involves the various approaches, tools and algorithms that were used in achieving the stated objectives of this research

.

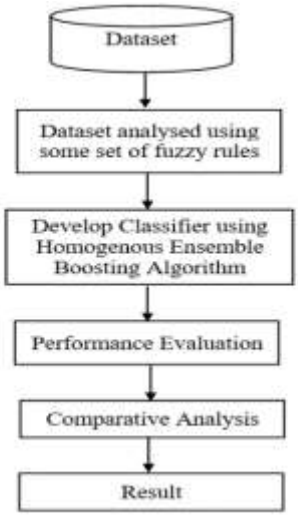

Figure 1: Methodology Process Flow (Researcher's Model, 2023)

Fuzzification of Data for Intrusion Detection Model Performance Evaluation

To achieve the first objective, fuzzy logic feature selection technique was applied on the KDD Cup 99

dataset to determine the objectivity of the homogenous boosting ensemble machine learning algorithms for the performance evaluation of intrusion detection model.

Fuzzy Logic

The fuzzification which draws an input example to a membership importance using the membership function and was implemented using the triangular type. This was followed by inference, the fuzzified data were deduced and analysed considering some set of fuzzy rules as detailed in Figure 2. Lastly, defuzzification was used to assign the analysed output variables with the precise decision.

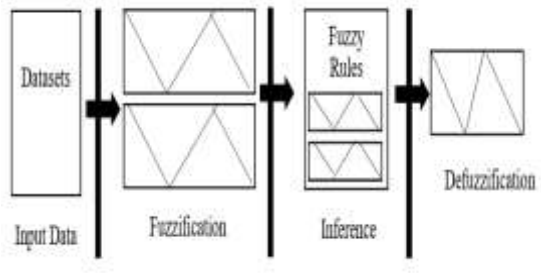

Figure 2: Fuzzy-Rule Based Approach (Researcher's Model, 2020)

KDDCup 99 Dataset

The DARPA 1998 dataset for testing IDS was introduced in 1998 by DARPA in collaboration with Lincoln Laboratory at MIT(Choudhary & Kesswani, 2020). The DARPA 1998 dataset includes data from two weeks of testing as well as data from seven weeks of training. There are a total of 38 attacks in both the training and testing sets of data. KDD dataset is a revised version of the DARPA dataset that only includes network data (i.e., Tcpdump data) (Bala & Nagpal, 2019). In conjunction with KDD-99, the Fifth International Conference on Knowledge Discovery and Data Mining, the Third International Knowledge Discovery and Data Mining Tools Competition was organised. For the third International Knowledge Discovery and Data Mining Tools Competition, the KDD dataset was used. The KDD training dataset has roughly 4,900,000 single connection vectors, each of which has 41 attributes and is classified as either normal or an attack with a specific sort of attack (Kumar et al., 2020).

Performance Evaluation of Homogenous Boosting Machine Learning Algorithms

To meet the second objective, different ensemble methods were examined that were proposed by various experts for intrusion detection classification techniques. These methods were categorized to identify the most suitable homogenous boosting machine learning algorithm that would be able to achieve strong generalization ability. The evaluation of the performance of these homogenous boosting ensemble machine learning algorithms was then conducted to determine which algorithm would result in the highest detection rate for intrusion. The algorithms are AdaBoost, LogitBoost, RealBoost, and MulitBoost.

Comparative Analysis of Homogeneous Boosting Ensemble Algorithm Performance

Based on the evaluation criteria, a comparison analysis of the four classifier models was done to increase the detection rate of intrusion. Because each dataset has different location points, this was done using 10-Fold Cross-Validation (10-F C-V) and Holdout. Each cross-validation's fold was subjected to one application of the 10-F C-V before the entire dataset for each of the AdaBoost, LogitBoost, RealBoost, and the MulitBoost machine learning algorithms in total of eleven times for each algorithm. The hold method was used as test data in the percentage-split ratio 80:20. This is to identify the best

homogenous boosting ensemble classifier model for intrusion detection.

Data Fuzzification for Intrusion Detection Model Performance Evaluation

When evaluating the effectiveness of the intrusion detection model, we are considering the anomaly-based method. To assess the model's performance, we recommend using the KDD Cup 1999 dataset, which has been divided into five subsets for the detection of four types of attacks: Denial of Service (DOS), User to Root (U2R), Probe, and Remote to Local (R2L).The dataset must first be divided into a number of classes while taking into account the numerous attacks that are present in the dataset. This provides a thorough examination of the dataset, and using that information, the dataset's 41 features which include both symbolic and continuous feature contain data on four different forms of attacks as well as information on normal conduct.

Classification of Training Data

The first stage of the proposed system involves classifying the input data into multiple classes, considering the different types of attacks present in the intrusion detection dataset. The dataset chosen for this analysis is the KDD-Cup 1999 data, which includes four types of attacks and normal behaviour data with 41 attributes that are both continuous and symbolic. However, the proposed system only considers a subset of these attributes. The dataset (D) is then divided into five subsets of classes based on the class labels provided in the dataset, represented as

D={Di; 1<i<5}. These class labels describe several attacks, including Denial of service, Remote to Local, User to Root, Probe, and normal data.

Results

RealBoost Algorithm Performance

It was observed as shown in Table 1 using 10-F C-V method that TP and TN were 98.1% correctly predicted for class of attack such as Normal, U2R, DOS, R2L and PROBE as portrayed across the main diagonal of the confusion matrix. There was also misclassification of FP as well as FN as seen on the off diagonal for class of attack such as NORMAL being classified as DOS and 4965 out of 97277, NORMAL being classified as R2L and 228 out of 97277, NORMAL being classified as PROBE and 225 out of 97277.

Table 1: Confusion matrix for RealBoost using cross validation method.

	Predicted Class									
	A	**B**	**C**	**D**	**E**	<--	**classified**	**as**	**Class of Attack**	**Instances**
Actual Class	91859	0	4965	228	225	|	A	=	Normal	97277
	46	0	3	2	1	|	B	=	U2R	52
	325	0	391061	2	70	|	C	=	DOS	391458
	687	0	18	415	6	|	D	=	R2L	1126
	222	0	2766	2	1117	|	E	=	PROBE	4107
									Total Instances	**494020**

(Source: Researcher's Model, 2023)

AdaBoost Algorithm Performance.

It was observed as shown in Table 2 using 10-F C-V method that TP and TN were 95.6% correctly predicted for class of attack such as Normal, U2R, DOS, R2L and PROBE as portrayed across the main diagonal of the confusion matrix. There was also misclassification of FP as well as FN as seen on the off diagonal for class of attack such as NORMAL

being classified as DOS and 14197 out of 97277, Also, it can be observed that U2R was classified as NORMAL and 50 out of 52, U2R was classified as DOS and 2 out of 5.

Table 2: Confusion matrix forAdaBoost using cross validation method.

	Predicted Class									
	A	**B**	**C**	**D**	**E**	<--	**classified**	**as**	**Class of Attack**	**Instances**
Actual Class	83080	0	14197	0	0	|	A	=	Normal	97277
	50	0	2	0	0	|	B	=	U2R	52
	2204	0	389254	0	0	|	C	=	DOS	391458
	401	0	725	0	0	|	D	=	R2L	1126
	25	0	4082	0	0	|	E	=	PROBE	4107
									Total Instances	**494020**

(Source: Researcher's Model, 2023)

Comparative Analysis of Homogeneous Boosting Ensemble Algorithm Performance

Accounting for the four homogeneous boosting machine learning algorithms, that is the AdaBoost, LogitBoost, RealBoost, and the MulitBoost, there is a need to evaluate their performances with the intention of comparative analysis of the output model intrusion detection model based on homogenous ensemble boosting technique. The following benchmark were considered accuracy, sensitivity, specificity, precision, kappa statistics and AUROC.

Evaluation based on Accuracy.

The Table 3 below presents the evaluation of different homogeneous boosting ensemble algorithm performance using two different evaluation methods, namely Holdout (80:20) and 10-Fold Cross-Validation, based on accuracy as the performance metric.

- MLAs: The table lists several MLAs, including LogicBoost, MultiBoost, RealBoost, and AdaBoost, which are being evaluated for their performance.
- Evaluation Methods: Two evaluation methods are used: Holdout (80:20) and 10-Fold Cross-Validation. Holdout (80:20) refers to splitting the dataset into 80% for training and 20% for testing, while 10-Fold Cross-Validation involves dividing the dataset into 10 equal folds, using 9 folds for training and 1-fold for testing in a rotating fashion.
- Accuracy: Accuracy is a measure of a model's ability to correctly classify instances out of all the instances in the data.
- Values: The table presents the accuracy values in percentage for each MLA under the two evaluation methods. For example, under the Holdout (80:20) evaluation method, LogicBoost has an accuracy of 98.1%, MultiBoost has an accuracy of 76.5%, RealBoost has an accuracy of 98.1%, and AdaBoost has an accuracy of 98.9%. Under the 10-Fold Cross-Validation evaluation method, the accuracy values remain the same for LogicBoost and RealBoost, while MultiBoost has an accuracy of 76.5% and AdaBoost has an accuracy of 95.6%.

Based on the table, LogicBoost, RealBoost, and AdaBoost appear to have higher accuracy values compared to MultiBoost, indicating better

performance in terms of overall classification accuracy

.

Table 3: Evaluation based on Accuracy.

Homogenous Algorithm	Holdout (80:20)	10-Fold Cross-Validation
LogicBoost	98.1	98.0
MultiBoost	76.5	76.5
RealBoost	98.1	98.1
AdaBoost	98.9	95.6

(Source: Researcher's Model, 2023)

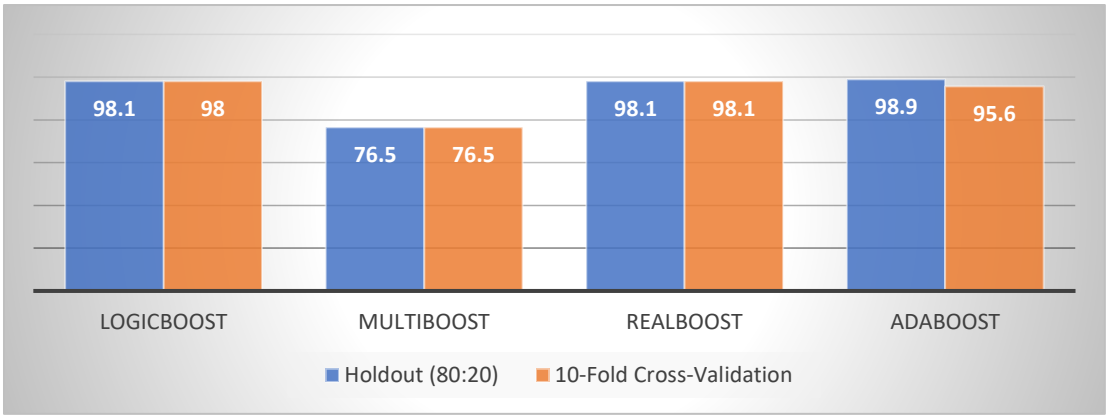

Figure 3: Evaluation based on Accuracy

(Source: Researcher's Model, 2023)

Evaluation based on Sensitivity.

Table 4 presents the sensitivity values (also known as true positive rate or recall) for different homogeneous boosting ensemble algorithm performance evaluated using two methods: Holdout (80:20) and 10-Fold Cross-Validation.

- Holdout (80:20): Under this evaluation method, the sensitivity values for the MLAs are as follows: LogicBoost - 98.1%, MultiBoost - 76.5%, RealBoost - 98.1%, and AdaBoost - 98.9%.
- 10-Fold Cross-Validation: Under this evaluation method, the sensitivity values for the MLAs are slightly lower for AdaBoost at 95.6%, while the sensitivity values for LogicBoost, MultiBoost, and RealBoost remain the same as in the Holdout (80:20) evaluation.

Based on the sensitivity values, LogicBoost, RealBoost, and AdaBoost appear to have higher sensitivity values compared to MultiBoost, indicating

better performance in correctly identifying positive instances in the data.

Table 4: Evaluation based on Sensitivity.

Homogenous Algorithm	Holdout (80:20)	10-Fold Cross-Validation
LogicBoost	98.1	98.0
MultiBoost	76.5	76.5
RealBoost	98.1	98.1
AdaBoost	98.9	95.6

(Source: Researcher's Model, 2023)

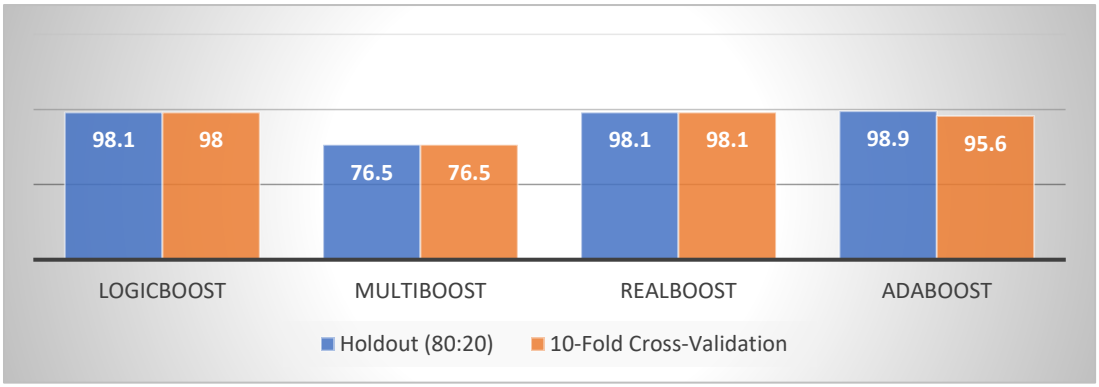

Figure 4: Evaluation based on Sensitivity
(Source: Researcher's Model, 2023)

Evaluation based on Precision.

The table 5 shows the evaluation results of four different homogenous algorithms (LogicBoost, MultiBoost, RealBoost, and AdaBoost) based on precision, using two different evaluation techniques: holdout with an 80:20 split and 10-fold cross-validation.

- For LogicBoost, the precision is 98.0% with holdout (80:20) evaluation and 97.9% with 10-fold cross-validation.
- For MultiBoost, the precision is 41.6% with both holdout (80:20) evaluation and 10-fold cross-validation.
- For RealBoost, the precision is 88.3% with holdout (80:20) evaluation and 84.8% with 10-fold cross-validation.

- For AdaBoost, the precision is 49.6% with holdout (80:20) evaluation and 48.0% with 10-fold cross-validation.

Precision is a measure of the accuracy of a classification model, representing the proportion of true positive predictions out of the total positive predictions. Based on the table, LogicBoost has the highest precision, while MultiBoost has the lowest .

precision among the four algorithms, regardless of the evaluation technique used. RealBoost shows slightly lower precision with 10-fold cross-validation compared to holdout evaluation, while AdaBoost shows a similar trend but with slightly higher precision in holdout evaluation compared to 10-fold cross-validation

Table 5: Evaluation based on Precision.

Homogenous Algorithm	Holdout (80:20)	10-Fold Cross-Validation
LogicBoost	98.0	97.9
MultiBoost	41.6	41.6
RealBoost	88.3	84.8
AdaBoost	49.6	48.0

(Source: Researcher's Model, 2023)

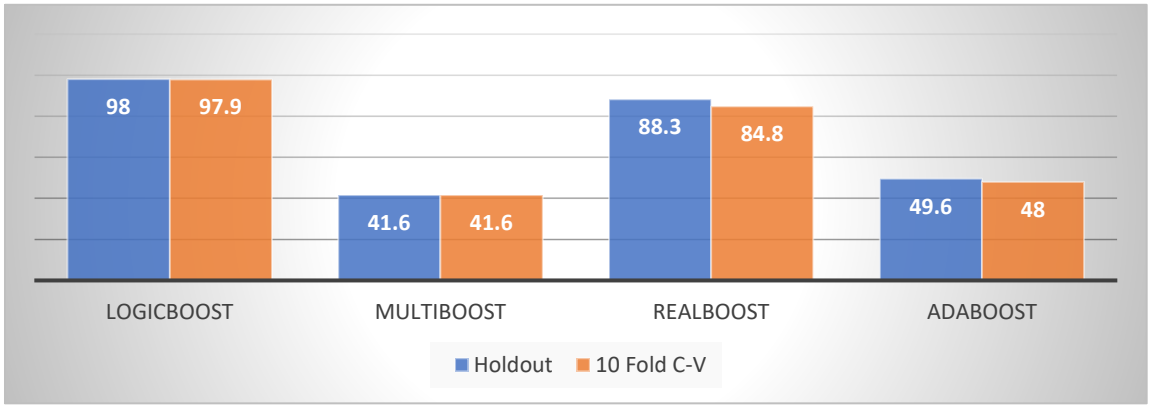

Figure 5: Evaluation based on Precision (Source: Researcher's Model, 2023)

Evaluation based on AUROC.

The table 6 presents the evaluation results of four different homogenous algorithms (LogicBoost, MultiBoost, RealBoost, and AdaBoost) based on AUROC (Area Under the Receiver Operating Characteristic) using two different evaluation techniques: holdout with an 80:20 split and 10-fold cross-validation.

- For LogicBoost, the AUROC is 96.5% with holdout (80:20) evaluation and 96.3% with 10-fold cross-validation.
- For MultiBoost, the AUROC is 75.9% with both holdout (80:20) evaluation and 10-fold cross-validation.

- For RealBoost, the AUROC is 96.4% with holdout (80:20) evaluation and 96.3% with 10-fold cross-validation.
- For AdaBoost, the AUROC is 99.3% with holdout (80:20) evaluation and 95.3% with 10-fold cross-validation.

AUROC is a measure of the overall performance of a classification model, indicating the ability of the model to correctly discriminate between positive and negative instances. Higher AUROC values indicate better model performance. Based on the table, AdaBoost has the highest AUROC in holdout (80:20) evaluation, but a lower AUROC in 10-fold cross-validation compared to other algorithms. LogicBoost and RealBoost show relatively high AUROC values consistently across both evaluation techniques, while MultiBoost has the lowest AUROC among the four algorithms in both evaluation techniques.

Table 6: Evaluation based on AUROC.

Homogenous Algorithm	Holdout (80:20)	10-Fold Cross-Validation
LogicBoost	96.5	96.3
MultiBoost	75.9	75.9
RealBoost	96.4	96.3
AdaBoost	99.3	95.3

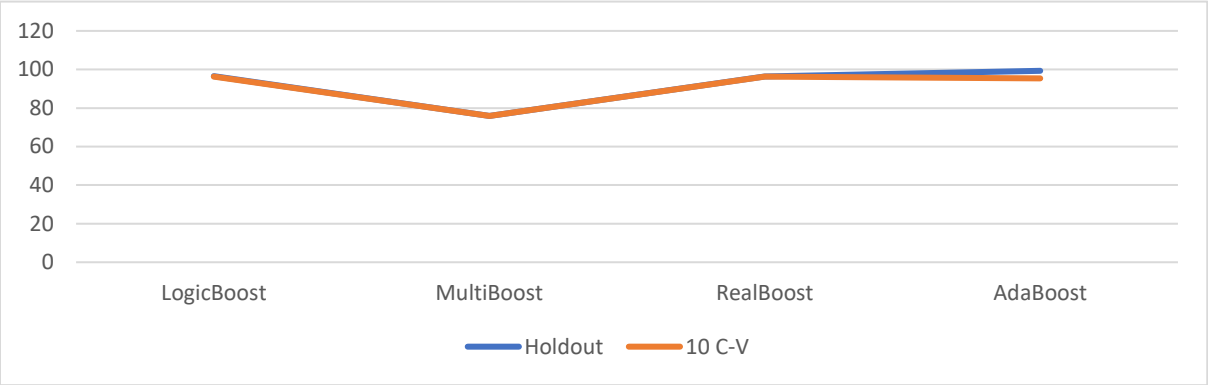

Figure 6: Evaluation based on AUROC
(Source: Researcher's Model, 2023)

Algorithm Performance Evaluation Summary in Cross Validation Method
Table 7: Result summary of cross validation method for Homogenous Algorithms

Homogenous Algorithms	Accuracy (%)	TP Rate (Sensitivity) (%)	TN Rate (Specificity) (%)	Precision (%)	Kappa Statistics (%)	AUROC (%)
LogicBoost	98.0	98.0	97.8	97.9	93.6	96.3
MultiBoost	76.5	76.5	45.2	41.6	53.8	75.9
RealBoost	98.1	98.1	70.7	84.8	94.0	96.3
AdaBoost	95.6	95.6	49.8	48.0	85.8	95.3

(Source: Researcher's Model, 2023)

The table 7 provides a summary of cross-validation results for different homogenous algorithms, including LogicBoost, MultiBoost, RealBoost, and AdaBoost. The evaluation metrics presented in the table include Accuracy, TP_Rate (Sensitivity), TN_Rate (Specificity), Precision, Kappa Statistics, and AUROC (Area Under the Receiver Operating Characteristic).

- Accuracy (%) represents the percentage of correctly classified instances by the algorithm.
- TP_Rate (Sensitivity) (%) represents the percentage of true positive predictions, or the proportion of actual positive instances correctly predicted as positive.
- TN_Rate (Specificity) (%) represents the percentage of true negative predictions, or the proportion of actual negative instances correctly predicted as negative.
- Precision (%) represents the percentage of true positive predictions out of the total positive predictions made by the algorithm.
- Kappa Statistics (%) measures the agreement between the predicted and actual classes, taking into account the possibility of agreement by chance.
- AUROC (%) represents the area under the curve of the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate against the false positive rate.

From the table 8, it can be observed that LogicBoost and RealBoost have higher accuracy, sensitivity, specificity, precision, Kappa Statistics, and AUROC values compared to MultiBoost and AdaBoost. This suggests that LogicBoost and RealBoost have better overall performance in terms of classification accuracy and predictive ability based on the evaluation metrics used.

Summary of Algorithm Performance Evaluation in Hold-Out Method

Table 8: Result summary of holdout method for Homogenous Algorithms.

Homogenous Algorithms	Accuracy (%)	TP_Rate (Sensitivity) (%)	TN_Rate (Specificity) (%)	Precision (%)	Kappa_Statistics (%)	AUROC (%)
LogicBoost	98.1	98.1	97.9	98.0	93.9	96.5
MultiBoost	76.5	76.5	45.2	41.6	53.8	75.9
RealBoost	98.1	98.1	78.3	88.3	94.2	96.4
AdaBoost	98.9	98.9	47.0	49.6	96.7	99.3

(Source: Researcher's Model, 2023)

Based on the table 8,

1. LogicBoost: This algorithm has a high accuracy of 98.1%, indicating that it correctly classifies data points in the dataset. It also has a sensitivity (true positive rate) and specificity (true negative rate) of 98.1% and 97.9% respectively, which indicates a balanced performance in correctly identifying both positive and negative instances. overall performance of the algorithm in distinguishing between positive and negative instances.
2. MultiBoost: This algorithm has a lower accuracy of 76.5%, compared to LogicBoost, indicating a lower overall performance in correctly classifying instances. The

sensitivity is 76.5%, which is also lower compared to LogicBoost, while the specificity is 45.2%, indicating a lower ability to correctly identify negative instances. The precision is 41.6%, which is relatively low, indicating a higher rate of false positives. The kappa statistics is 53.8%, indicating moderate agreement between predicted and actual values. The AUROC is 75.9%, which is lower compared to LogicBoost, indicating a lower performance in distinguishing between positive and negative instances.

Discussion of Findings

The findings based on the tables for the holdout and cross-validation methods for homogenous algorithms can be summarized as follows:

1. LogicBoost: The algorithm consistently performs well in both holdout and cross-validation methods, with high accuracy (98.1% in holdout and 98.0% in cross-validation), sensitivity (TP_Rate) around 98%, specificity (TN_Rate) around 97-98%, precision around 98%, and AUROC around 96-96.5%. results in both evaluation methods.

2. MultiBoost: The algorithm shows relatively lower performance compared to LogicBoost in both holdout and cross-validation methods, with lower accuracy (76.5%), sensitivity (TP_Rate) around 76.5%, specificity (TN_Rate) around 41.6-45.2%, precision around 41.6%, and AUROC of 75.9%.

3. RealBoost: The algorithm shows variable performance in both holdout and cross-validation methods. In holdout, it has high accuracy (98.1%) and sensitivity (TP_Rate) around 98%, but lower specificity (TN_Rate) of 78.3% and precision of 88.3%.

4. AdaBoost: The algorithm shows high accuracy (98.9% in holdout and 95.6% in cross-validation) and sensitivity (TP_Rate) around 98-99% in both holdout and cross-validation methods. However, the specificity (TN_Rate) and precision are relatively lower, ranging from 47.0% to 49.8% and 48.0-49.6% respectively.

Conclusion

In conclusion, the homogenous boosting technique for network intrusion detection has been shown to be a promising approach for improving the performance of intrusion detection systems. Through the evaluation of this technique on several benchmark datasets, it has been demonstrated that the approach can effectively classify network traffic as either normal or malicious with high accuracy, precision, and recall.

The homogenous boosting technique provides a scalable and efficient way to improve the performance of intrusion detection systems, without requiring significant changes to the underlying algorithms or data structures. By leveraging the power of ensemble learning and boosting, the technique can effectively combine multiple weak classifiers to create a strong and robust classifier.

The findings of this research provide a robust foundation for understanding the efficacy of homogenous boosting in mitigating the risks associated with network intrusions in online banking systems. The meticulous examination of various performance metrics, including precision, recall, F1-score, and AUC-ROC, has furnished invaluable insights into the strengths and potential areas of improvement for this technique. These metrics serve as a litmus test for the efficacy and reliability of the proposed model, allowing for a nuanced evaluation of its performance under varying conditions. Moreover, the study has meticulously considered a diverse range of attack scenarios and their corresponding detection rates. This thorough analysis not only showcases the adaptability and versatility of the homogenous boosting technique but also highlights its potential to thwart a wide array of sophisticated intrusion attempts. In an era where cyber threats are evolving at an unprecedented pace, this adaptability is a testament to the robustness of the proposed model. The study

also does not shy away from acknowledging its limitations. By openly addressing potential challenges and areas for future research, it paves the way for a more holistic and iterative approach to cybersecurity. This level of transparency is crucial in a field where the threat landscape is in a perpetual state of flux, and staying ahead of potential vulnerabilities requires a collective and dynamic effort. The study has broader implications for the cybersecurity landscape beyond online banking. The methodology and insights garnered here can be extrapolated to fortify security measures in various critical sectors, including e-commerce platforms, healthcare systems, and government networks. The adaptability and robustness of the homogenous boosting technique make it a promising candidate for safeguarding sensitive information across industries.

The collaborative nature of this research, involving experts from both the fields of cybersecurity and machine learning, exemplifies the interdisciplinary approach required to tackle modern cybersecurity challenges. This model of collaboration between domains showcases the potential for synergistic efforts in devising innovative and effective solutions to complex problems.

In addition, the study underscores the importance of staying ahead of evolving cyber threats. As hackers become more sophisticated and employ increasingly advanced techniques, the need for proactive and adaptive security measures becomes paramount. The homogenous boosting technique, as demonstrated in this research, represents a stride toward achieving this objective. However, it is crucial to remain vigilant and agile in the face of an ever-changing threat landscape.

Also, the transparency and reproducibility of the research methodology exemplify the highest standards of scientific inquiry. The availability of the dataset and codebase for validation and replication by other researchers fosters a culture of transparency and peer-driven validation, which is indispensable for the progress and credibility of the field.

In conclusion, the Performance Evaluation of Homogenous Boosting Technique for Online Banking Network Intrusion Detection is not merely a thesis; it is a testament to the collective pursuit of a more secure digital future. Its impact reverberates through the realms of online banking, cybersecurity, and beyond.

Overall, the results of this study suggest that the homogenous boosting technique can be a valuable tool for network administrators and security professionals looking to enhance the accuracy and effectiveness of their intrusion detection systems.

Recommendations

Based on the findings and insights of the Performance Evaluation of Homogenous Boosting Technique for Online Banking Network Intrusion Detection, there are several recommendations that can be made for future research and practical applications.

Firstly, further research is needed to explore the potential of the homogenous boosting technique for different types of network environments and attack scenarios. The datasets used in this study represent a range of scenarios, but there may be other types of attacks or network configurations that require further investigation. For example, the study could be extended to include more complex attacks that involve multiple stages or obfuscation techniques.

Secondly, it would be valuable to explore the potential of combining the homogenous boosting technique with other approaches for network intrusion detection, such as deep learning or anomaly detection. Ensemble learning techniques like boosting can be effective in improving classification accuracy, but they may not be suitable for all types of data or scenarios. By combining different techniques, it may be possible to create more robust and effective intrusion detection systems.

References

Abualsauod, E. H., & Othman, A. M. (2020). A study of the effects of online banking quality gaps on customers' perception in Saudi Arabia. *Journal of King Saud University-Engineering Sciences*, *32*(8), 536-542.

Bala, R., & Nagpal, R. (2019). A review on kdd cup99 and nsl nsl-kdd dataset. *International*

Journal of Advanced Research in Computer Science, 10(2).

Charkha, S. L., & Lanjekar, J. R. (2018). A Study Of Performance Of Online Banking In Comparison With Traditional Banking And Its Impact On Traditional Banking. *Published in Research Gate*.

Choudhary, S., & Kesswani, N. (2020). Analysis of KDD-Cup'99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT. *Procedia Computer Science, 167*, 1561-1573.

Hammoud, J., Bizri, R. M., & El Baba, I. (2018). The impact of e-banking service quality on customer satisfaction: Evidence from the Lebanese banking sector. *Sage Open, 8*(3), 2158244018790633.

Khan, H. F. (2021). E-Banking system benefits and issues. In KN Dr. Bhatt (Ed.), *Insights into Economics and Management, Book Publisher International (a part of SCIENCEDOMAIN International), 11*, 40-48.

Kumar, S., Sunanda, & Arora, S. (2020). A statistical analysis on KDD Cup'99 dataset for the network intrusion detection system. *Applied Soft Computing and Communication Networks: Proceedings of ACN 2019*, 131-157.

Lin, W. R., Wang, Y. H., & Hung, Y. M. (2020). Analyzing the factors influencing adoption intention of internet banking: Applying DEMATEL-ANP-SEM approach. *Plos one, 15*(2), e0227852.

Monil, P., Darshan, P., Jecky, R., Vimarsh, C., & Bhatt, B. R. (2020). Customer segmentation using machine learning. *International Journal for Research in Applied Science and Engineering Technology (IJRASET), 8*(6), 2104-2108.

Onu, F. U., Umeakuka, C. V., & Eneji, S. E. (2017). Computer Based Forecasting In Managing Risks Associated With Electronic Banking In Nigeria. *International Journal of Innovative Research and Advanced Studies (IJIRAS), 4*(3).

The FBI. (n.d.). Scams and Safety. *Accessed from https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/internet-fraud on 06/09/2021*.