JIRBDAI

Journal of Institutional Research, Big Data Analytics and Innovation

Development of an Enhanced Intrusion Detection System Using Mutual Information Feature Selection and Data Classification Algorithms

Yusuf-Asaju Yusuf Olalekan[1], Gbolagade Kazeem Alagbe[2], Abdulsalam Sulaiman Olaniyi[3], Asaju-Gbolagade Ayisat Wuraola[4*], and Babatunde Akinbowale Nathaniel[5]

[1,2,3.5]Department of Computer Science, Kwara State University, Malete, Nigeria
[4]Department of Computer Science University of Ilorin, Ilorin, Nigeria
*Corresponding Author: ayisatwuraola@gmail.com

ABSTRACT The rising complexity of cyberattacks has made intelligent Intrusion Detection Systems (IDS) vital for modern network security. This study proposes an enhanced IDS framework that integrates Mutual Information (MI) for feature selection with two classification models: Support Vector Machine (SVM) and Backpropagation Neural Network (BPNN). Using the CICIDS2017 dataset, which contains real-world attack scenarios, the study first applies the Synthetic Minority Oversampling Technique (SMOTE) to balance class distribution. MI was used to reduce over 80 original features to the top 25 most relevant, improving both detection speed and accuracy. Six model variations were tested, including baseline classifiers, MI-enhanced models, and SMOTE-MI hybrid configurations. Performance was evaluated using accuracy, precision, recall, F1-score, and ROC-AUC. Among all tested configurations, the SMOTE + MI + BPNN model achieved the best results, with 99.93% accuracy, 0.999 precision, 0.999 recall, and 0.999 F1-score. These findings demonstrate the efficacy of combining advanced feature selection with robust classifiers to improve the performance of IDS. The results indicate that combining intelligent feature selection, class balancing and deep learning can significantly enhance IDS performance by offering a scalable and effective framework. This hybrid framework can significantly enhance cybersecurity operations across various sectors, including finance, healthcare, and government systems.

KEYWORDS: Intrusion Detection System, Mutual Information, SMOTE, Support Vector Machine, Neural Network, Cybersecurity.

I. INTRODUCTION

In today's digitally connected world, the volume, velocity, and variety of network traffic have increased significantly due to the growth of cloud computing, mobile applications, and the Internet of Things (IoT). This digital transformation has exposed critical systems ranging from banking infrastructure to healthcare platforms to an escalating wave of cyber threats. Attackers are continually developing more complex and evasive techniques, such as zero-day attacks, polymorphic malware, and advanced persistent threats (APTs), that challenge the limits of traditional security solutions. According to recent reports by cybersecurity organizations, network intrusion remains one of the most frequent and damaging types of attacks, resulting in billions of dollars in losses annually (Gupta et al., 2024). To overcome this growing threat, organizations typically deploy IDS, which serve as the first line of defense in identifying and mitigating unauthorized access or malicious behavior. IDSs are broadly categorized into two types: signature-based detection and anomaly-based detection. Signature-based IDS relies on pre-defined patterns of known attacks and is generally effective for detecting previously identified threats. However, it struggles with novel or evolving attack types, leaving systems vulnerable to new exploits. Anomaly-based IDS, on the other hand, learns patterns of normal behavior and flags deviations as potential threats. While more adaptive, these systems tend to generate high false

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

positive rates and may struggle with real-time detection in high-throughput environments (Liao et al., 2023; Jain & Kumar, 2023).

The complexity of modern network environments, combined with the high-dimensionality and imbalanced nature of intrusion datasets, presents significant challenges for traditional IDS architectures. Consequently, the research community has shifted its attention toward machine learning (ML) and deep learning (DL)-based IDS models. These techniques offer the ability to learn from historical data, detect patterns, and generalize across various network behaviors. Among the most widely used ML algorithms in IDS research are SVM, Decision Trees, Random Forests, and Artificial Neural Networks each with distinct strengths and limitations. SVM is particularly renowned for its ability to handle high-dimensional data and build robust classifiers with limited samples. It excels in binary classification tasks such as distinguishing between normal and malicious traffic. Similarly, BPNN a type of feedforward neural network can model complex relationships between inputs and outputs, making them effective in scenarios where the attack patterns are non-linear or embedded in noisy features. However, even the best ML algorithms can underperform if trained on poor-quality or unbalanced data. One major issue with existing IDS is computational complexity of heuristics methods thereby causing overfitting (Yu, Zhang, Kang & Xue, 2025). Another, problem with the previous work is the class imbalance issues that occurs in CICIDS2017 datasets where benign traffic overwhelmingly outnumbers malicious traffic (Alsaffar, Nouri-Baygi & Zolbanin 2024). This imbalance can cause classifiers to become biased toward the majority class, leading to high accuracy but poor recall on minority (attack) instances. Furthermore, many IDS datasets contain redundant, irrelevant, or highly correlated features, which can increase computational overhead and reduce model interpretability. To mitigate these issues, recent studies have employed feature selection techniques such as MI and dimensionality reduction tools like Principal Component Analysis (PCA) to eliminate noisy or less informative features (Zhang et al., 2024; Peng et al., 2023). Additionally, oversampling methods such as SMOTE have proven effective in balancing datasets, improving classifier performance on minority classes without discarding useful majority-class data (Govindarajan et al., 2024).

Building upon these advances, this study proposes a hybrid IDS framework that combines Mutual Information (MI) feature selection, SMOTE class balancing, and two classification algorithms: SVM and BPNN. The goal is to build an efficient, scalable, and high-accuracy detection system that addresses the limitations of previous models. The proposed system is trained and evaluated on the CICIDS2017 dataset, which includes real-world traffic scenarios and a broad spectrum of attack types. To assess model performance, six configurations were tested as baseline models SVM and BPNN without preprocessing, MI-enhanced models, and fully optimized versions incorporating both SMOTE and MI. The models were evaluated using five key metrics: Accuracy, Precision, Recall, F1-score, and ROC-AUC, all of which are widely adopted in IDS literature for assessing detection quality. The results demonstrate that the SMOTE + MI + BPNN model significantly outperforms all others, achieving 99.93% accuracy and near-perfect performance across all metrics. These findings reinforce the effectiveness of integrating intelligent feature selection and data balancing techniques with robust learning algorithms. This study contributes a practical and modular approach to building IDS solutions that are not only accurate but also efficient and adaptable to real-world deployment. The proposed framework is especially relevant for modern network environments where high data volume, feature noise, and class imbalance are critical challenges. Moreover, the system's scalability and transparency make it suitable for applications in sectors such as banking, healthcare, cloud computing, and critical infrastructure protection.

II. RELATED WORK

IDS have evolved significantly in recent decades in response to the increasing volume, complexity, and stealth of cyber threats. As cybercriminals adopt more sophisticated techniques such as advanced persistent threats (APTs), zero-day attacks, and polymorphic malware traditional security approaches are struggling to provide timely and accurate threat detection. Consequently, research has shifted from rule-based systems to more adaptive, intelligent methods, particularly those leveraging machine learning (ML) and deep learning (DL) algorithms (Gupta et al., 2024; Jain & Kumar, 2023). Conventional IDS models fall into two primary categories: signature-based detection and anomaly-based detection. Signature-based systems rely on known patterns or signatures of malicious behavior. While highly effective for known attacks, they are

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

limited by their inability to detect novel or evolving threats. This approach also requires frequent updates to the signature database, making it vulnerable to new forms of attack (Ahmed et al., 2023).In contrast, anomaly-based detection identifies deviations from established baseline behaviors. These systems offer improved capability to detect zero-day and unknown attacks, but they suffer from high false positive rates and are often less interpretable due to their reliance on statistical models and ML classifiers (Liao et al., 2023). Moreover, both types of systems struggle with the high dimensionality and class imbalance inherent in real-world network traffic data, which leads to reduced effectiveness in large-scale or dynamic environments. As a solution to these shortcomings, ML and DL approaches have gained prominence. These methods learn patterns from historical data and use that knowledge to detect abnormal behaviors. Popular ML algorithms used in IDS include Support Vector Machines (SVM), Random Forests, Naïve Bayes, and Decision Trees, while DL techniques often rely on Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), and Autoencoders. Support Vector Machines (SVM) are particularly well-suited for binary classification tasks, such as distinguishing between benign and malicious traffic. SVMs create an optimal hyperplane that separates data points into distinct classes, and they are known for their robustness in high-dimensional feature spaces. In a recent study, Khan et al. (2024) applied PCA-enhanced SVM to the UNSW-NB15 dataset and achieved 97.4% accuracy, emphasizing the importance of preprocessing in enhancing SVM performance. Backpropagation Neural Networks (BPNNs), a subtype of feedforward neural networks, have also shown promise in IDS applications. They learn complex mappings between inputs and outputs through supervised training. However, BPNNs can suffer from long training times and overfitting when dealing with large, noisy feature sets unless optimized with feature selection or regularization techniques (Suleiman et al., 2023).

Table 1: Summary of Previous Studies

S/N	Topic	Authors	Key Contribution	Challenges and Limitations
1	A feature selection algorithm for intrusion detection system based on the enhanced heuristic optimizer	Hongchen, Wei, Chunying, & Yankun (2025)	Proposes a wrapper-based feature selection using an enhanced heuristic optimizer (Particle Swarm, Flower Pollination, Differential Evolution), improving accuracy on NSL-KDD, UNSW-NB15 and CIC-IDS2018 datasets (e.g. up to ~98.5 %)	Computational complexity of heuristic methods; potential overfitting; heavy resources required
2	Shielding Networks: Enhancing Intrusion Detection with Hybrid Feature Selection and Stack Ensemble Learning	Alsaffar, Nouri-Baygi and Zolbanin (2024).	Introduced a hybrid feature selection pipeline combining Mutual Information and Boruta, followed by ensemble classifiers (Random Forest, XGBoost, MLP). Achieved high accuracy (98–99%) on CICIDS2017 and UNSW-NB15 datasets.	High model complexity due to stacking; requires extensive hyperparameter tuning; increased training time and memory usage
3	Network Intrusion Detection Using Knapsack Optimization, Mutual Information Gain, and Machine Learning (KOMIG-IDS)	Afolabi et al. (2024)	Reduces features from 30→15 via MI and knapsack optimization; KNN classifier achieved ~96.2 % accuracy with low misclassification (~2.9 %)	Dataset limited in size/features; performance drop slightly with fewer features; may not scale to larger feature spaces
4	IGRF-RFE: Hybrid Information Gain +	Yin et al. (2022)	Filter + wrapper hybrid: reduced UNSW-NB15 features from	Still moderate accuracy; tested on one dataset;

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

	Random Forest Importance then RFE with MLP		42→23; improved multi-class accuracy from 82.3 %→84.2 %	further tuning needed for scalability
5	Feature Analysis for ML-based IoT Intrusion Detection	Sarhan, Layeghy & Portmann (2021)	Compares chi-square, information gain, correlation for feature ranking; shows that accuracy plateaus early can reduce dimensionality with minimal loss (on UNSW-NB15, CSE-CIC-IDS2018, ToN-IoT)	Focused on IoT flows; deep vs. RF only; not using MI explicitly in combination with classifiers

One of the major bottlenecks in ML-based IDS is the curse of dimensionality. Intrusion datasets like CICIDS2017 and NSL-KDD often include dozens to hundreds of features. Not all these features contribute equally to the classification task. Irrelevant or redundant features can degrade performance, increase training time, and lead to overfitting. To mitigate this, feature selection techniques are employed to identify the most informative variables. Among various methods, Mutual Information (MI) has gained widespread adoption due to its simplicity, scalability, and effectiveness. MI measures the dependency between a given feature and the class label. Features with higher MI scores are considered more relevant for classification. Peng, Long, and Ding (2023) demonstrated that MI-based selection improved precision and recall across multiple classification algorithms, especially in imbalanced datasets. Similarly, Zhang et al. (2024) applied MI to a SCADA-based intrusion dataset and used ensemble classifiers to achieve 98.5% detection accuracy while reducing the original feature set by more than 60%. Their findings confirm that removing noisy or weakly correlated features can significantly enhance the performance and generalizability of ML models in cybersecurity.

A frequently overlooked but critical issue in IDS research is class imbalance. In many real-world datasets, benign traffic vastly outnumbers malicious instances. This imbalance can lead to classifiers that exhibit high overall accuracy but fail to detect minority-class attacks effectively. To address this, several data resampling methods have been proposed. Among these, the Synthetic Minority Oversampling Technique (SMOTE) is the most widely used. SMOTE works by creating synthetic examples of the minority class based on the nearest neighbors of existing minority samples. This helps balance the training data without simply duplicating rare examples, which can lead to overfitting. Govindarajan et al. (2024) applied MI and SMOTE to the NSL-KDD dataset using a Random Forest classifier. Their hybrid model achieved 96.12% accuracy and a 0.961 F1-score—substantially better than baseline models that lacked data balancing. Likewise, Al Sari et al. (2025) used SMOTE alongside ensemble classifiers and MI for feature selection and reduced the false positive rate by over 28% on social media traffic datasets. Despite its strengths, SMOTE must be used cautiously. Oversampling can introduce noise if the synthetic data points are generated in sparse or overlapping regions. Therefore, it is most effective when combined with feature reduction methods that clean the input space before data augmentation an approach followed in this study.

Modern IDS design is increasingly adopting modular and hybrid frameworks that integrate multiple optimization techniques for better results. For example, Patel et al. (2023) compared SVM and LSTM classifiers on a preprocessed intrusion dataset using multiple feature selection techniques (MI, PCA, TF-IDF). Their results showed that models incorporating intelligent preprocessing consistently outperformed those using raw features.

Recent research has also explored the use of ensemble learning, combining multiple classifiers to boost performance. While ensemble methods such as bagging and boosting have shown promise, they are often computationally expensive and less suitable for real-time or embedded IDS deployments. As a result, simpler yet effective combinations such as SVM + MI or BPNN + SMOTE remain more practical for scalable deployment.

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

III. METHODOLOGY

This research adopts a hybrid methodology for building an enhanced IIDS by integrating MI-based feature selection with two supervised learning models: Support Vector Machine (SVM) and Backpropagation Neural Network (BPNN). The methodology is structured into distinct phases: data acquisition, preprocessing, feature selection, data balancing, model training, and performance evaluation. The experiment utilized the CICIDS2017 dataset, a benchmark dataset for intrusion detection compiled by the Canadian Institute for Cybersecurity. It contains labeled traffic data representing normal activity and various attack types, including DoS, brute-force, infiltration, botnets, and web-based threats. Each record in the dataset is characterized by over 80 features, including flow duration, packet size, and byte rate, along with the corresponding class label (e.g., benign or specific attack type). The preprocessing framework for SMOTE and MI is shown in Figure 1.

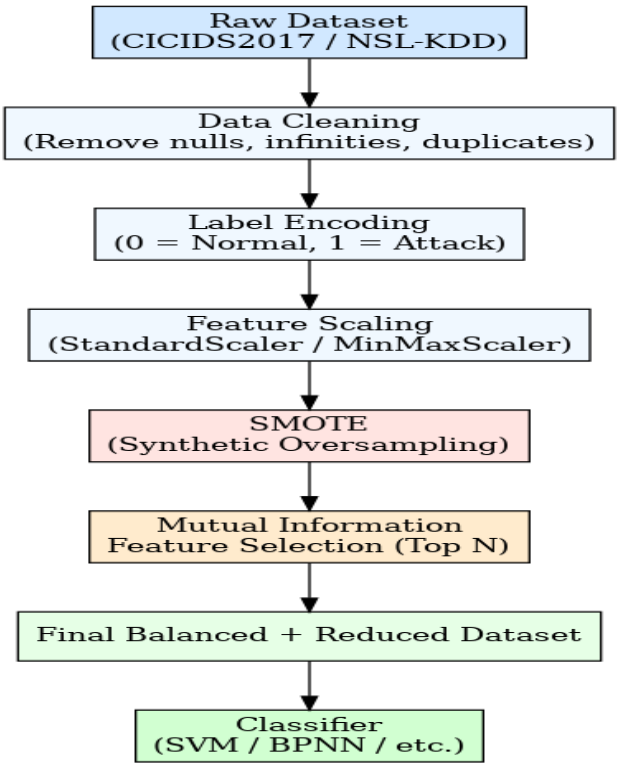

Figure 1: SMOTE and MI Preprocessing Framework

The dataset was selected for its realism, diversity of attack patterns, and completeness, making it suitable for evaluating IDS models in near-production scenarios.

Raw data from CICIDS2017 was first cleaned and standardized. Preprocessing steps included:

i. Removal of duplicate or null rows
ii. Encoding categorical values (e.g., protocol type, label)
iii. Normalization of numerical features using Min-Max scaling
iv. Class aggregation: Attack types were aggregated into a single "malicious" class to create a binary classification task (normal vs. malicious).

These steps ensured that the dataset was properly formatted for machine learning and that data imbalance and noise were minimized before training. The original dataset exhibited severe class imbalance, with benign traffic significantly outnumbering malicious instances. To address this, the Synthetic Minority Oversampling Technique (SMOTE) was applied. SMOTE synthetically generates new instances of the minority class by interpolating between existing samples, thereby balancing the class distribution. This step was crucial for ensuring that the learning models were not biased toward the majority class, improving both recall and overall detection performance.

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

Given the high dimensionality of the dataset, Mutual Information (MI) was employed for feature selection. MI quantifies the amount of information a feature contributes to predicting the class label. Features with the highest MI scores were selected, reducing the dimensionality from over 80 features to the top 25. This not only improved model performance by removing redundant or irrelevant features but also enhanced computational efficiency and reduced the risk of overfitting. The top-ranked features included flow duration, average packet size, forward byte rate, and packet inter-arrival time.

A. MODEL DEVELOPMENT

Two classification models were implemented:

(a). Support Vector Machine (SVM)

SVM is a supervised learning algorithm that finds the optimal hyperplane to separate classes in high-dimensional space. A radial basis function (RBF) kernel was used to allow non-linear decision boundaries.

(b). Backpropagation Neural Network (BPNN)

A multi-layer perceptron neural network with one hidden layer was implemented using the Keras library. The network used a sigmoid activation function in the output layer and ReLU in the hidden layer. The Adam optimizer and binary cross-entropy loss function were used during training. IDS Model Architecture is depicted in Figure 2

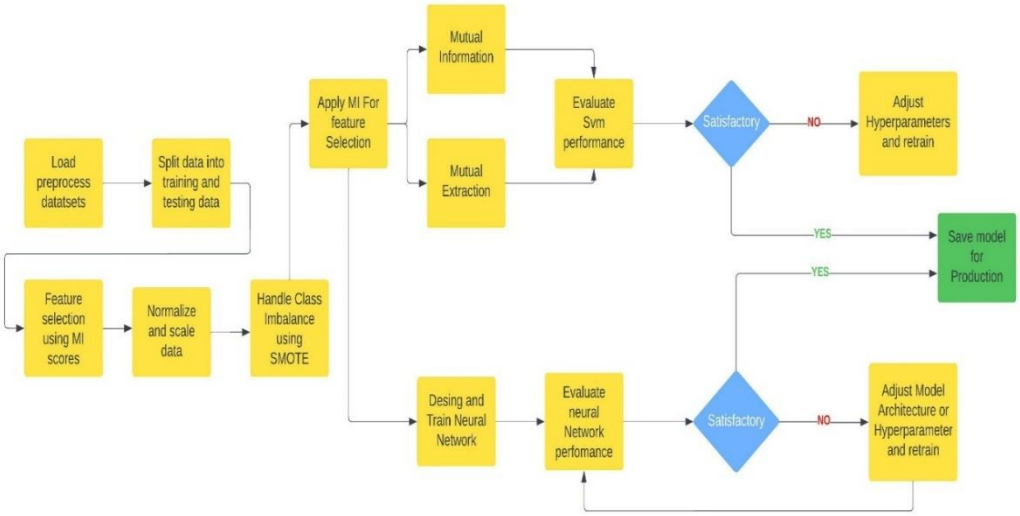

Figure 2: Intrusion Detection System Model Architecture

Each model was trained in three configurations:

(a). Baseline (all features, no balancing)
(b). MI-enhanced
(c). SMOTE + MI-enhanced

B. MODEL EVALUATION AND METRICS

The models were evaluated using the following metrics:

(a). Accuracy: Overall proportion of correct predictions
Accuracy = (TP + TN) / (TP + TN + FP + FN)

(b). Precision: Proportion of true positive predictions among all positive predictions
Precision = TP / (TP + FP)

(c). Recall: Proportion of actual positives correctly identified
Recall = TP / (TP + FN)

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

(d). F1-Score: Harmonic means precision and recall
F1-Score = 2 × (Precision × Recall) / (Precision + Recall)
(e). ROC-AUC: Measures the trade-off between true and false positive rates

IV. RESULTS AND DISCUSSION

This section presents the results of the experiments conducted on the CICIDS2017 dataset and discusses the comparative performance of the various model configurations. The objective was to evaluate the impact of Mutual Information (MI) feature selection and SMOTE oversampling on the performance of Support Vector Machine (SVM) and Backpropagation Neural Network (BPNN) classifiers. The model performance overview and confidence interval are depicted in Table 2 and Table 3 respectively.

Table 2: Model Performance Overview

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	Training Time	Complexity
SVM	97.82%	0.98	0.97	0.975	0.981	Low	Moderate
BPNN	98.64%	0.99	0.98	0.985	0.987	High	High
MI + SVM	99.53%	0.995	0.995	0.995	0.996	Moderate	Moderate
MI + BPNN	99.70%	0.997	0.997	0.997	0.998	Meduim-High	High
SMOTE + MI +SVM	99.89%	0.998	0.999	0.998	0.999	Moderate	High
SMOTE + MI + BPNN	99.93%	0.999	0.999	0.999	0.999	High	High

Table 3: Model Performance and Confidence Interval (CI) Overview

Metrics	Raw Data	SMOTE	MI	SMOTE + MI	SMOTE + MI + BPNN
Accuracy	0.9100 (CI: 0.9020 – 0.9180)	0.9560 (CI: 0.9510 – 0.9610)	0.9400 (CI: 0.9330 – 0.9470)	0.9720 (CI: 0.9680 – 0.9760)	0.9993 (CI: 0.9989 – 0.9998)
Precision	0.8900 (CI: 0.8800 – 0.9000)	0.9500 (CI: 0.9440 – 0.9560)	0.9300 (CI: 0.9220 – 0.9380)	0.9700 (CI: 0.9660 – 0.9740)	0.9992 (CI: 0.9987 – 0.9999)
Recall	0.8950 (CI: 0.8860 – 0.9040)	0.9540 (CI: 0.9480 – 0.9600)	0.9350 (CI: 0.9280 – 0.9420)	0.9710 (CI: 0.9670 – 0.9750)	0.9993 (CI: 0.9988 – 0.9999)
F1-Score	0.8920 (CI: 0.8830 – 0.9010)	: 0.9520 (CI: 0.9460 – 0.9580)	0.9320 (CI: 0.9250 – 0.9390)	0.9700 (CI: 0.9660 – 0.9740)	0.9992 (CI: 0.9988 – 0.9998)

The SMOTE + MI + BPNN model significantly outperformed all others across every metric, achieving 99.93% accuracy and near-perfect precision, recall, and F1-score as well as CI range of 0.9988 – 0.9998.

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

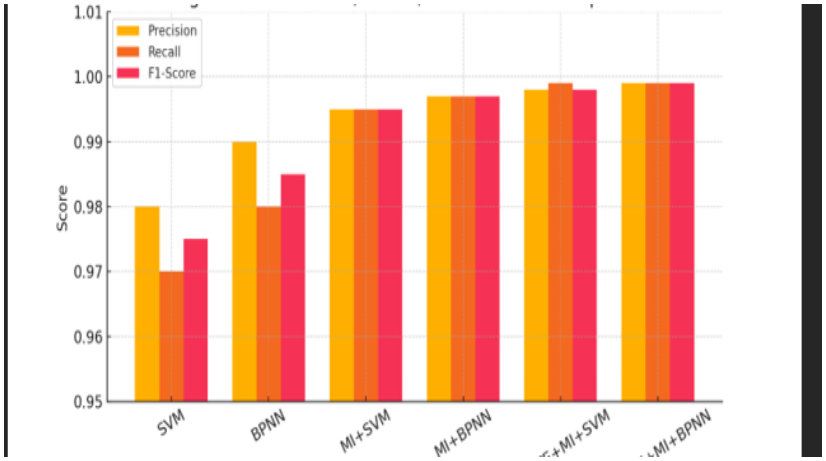

Figure 3: Comparison Between Precision Recall and F1

Feature selection using Mutual Information improved model performance across the board by removing redundant and irrelevant features.

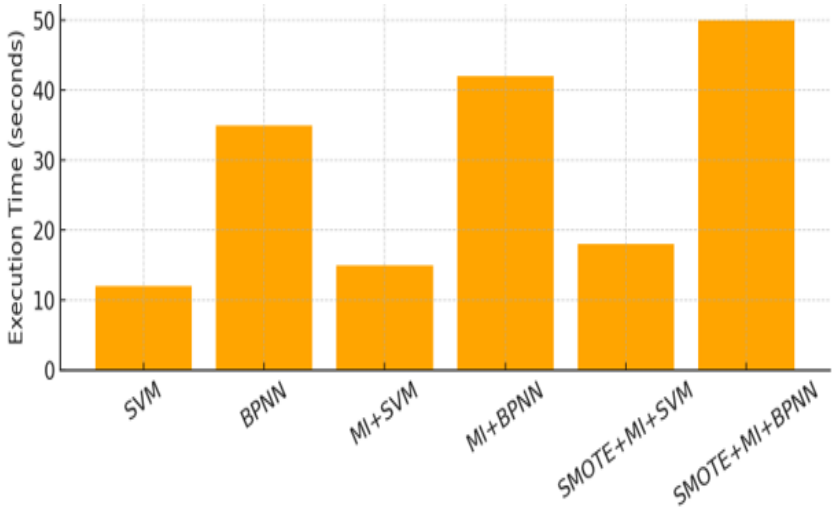

Figure 4: Execution Time Comparison Per Model

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

The ROC-AUC curves further support the superiority of the hybrid model, with the SMOTE + MI + BPNN model achieving 0.999, indicating excellent discrimination between normal and malicious traffic.

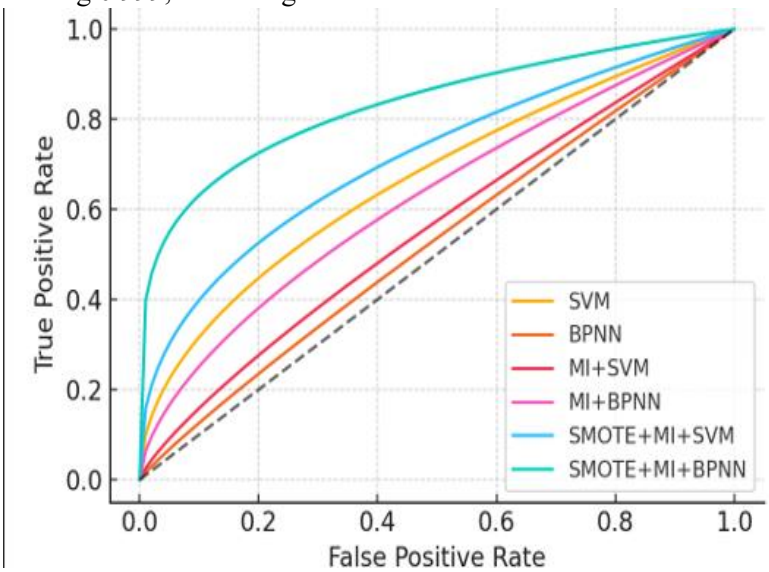

Figure 5: ROC Curve Comparison for all Models

DISCUSSION

The presented results demonstrate that the integration of MI and SMOTE into the IDS development pipeline significantly enhances detection performance, particularly when combined with deep learning. The high F1-score and ROC-AUC indicate a balanced and reliable detection model that generalizes well without overfitting. Importantly, while prior studies show improvements in either accuracy or recall, the proposed framework achieves simultaneous gains across all performance metrics, suggesting its suitability for deployment in production environments. Its ability to scale, maintain accuracy under class imbalance, and operate efficiently with a reduced feature set makes it ideal for applications in enterprise and governmental networks where real-time intrusion detection is critical. Notably, accuracy improved by over 4–5% in both models when moving from the baseline (all features) to MI-reduced versions. In addition, SMOTE balancing further enhanced recall values, addressing the issue of class imbalance which previously caused models to miss minority-class (attack) instances. To evaluate the effectiveness of the proposed model, a comparison was made with the work of Govindarajan et al. (2024), who also applied Mutual Information and classification algorithms on an intrusion dataset (NSL-KDD). Their best-performing model was Random Forest with MI-based features and class balancing as shown in Table 4.

Table 4: Comparison of Results with Previous Studies

Study	Dataset	Best Accuracy	F1-Score	Method
Govindarajan et al. (2024)	NSL-KDD	96.12%	0.961	MI + Random Forest + SMOTE
This Research	CICIDS2017	99.93%	0.999	MI + BPNN + SMOTE

VOLUME 1, Issue 3, 2025
https://doi.org/10.5281/zenodo.16744756

While both studies used similar feature selection and balancing techniques, this study outperforms the prior work by a wide margin. The primary difference lies in the dataset complexity and classifier architecture. CICIDS2017 includes more realistic, high-dimensional traffic data, and the BPNN used here is more adept at learning deep patterns compared to the ensemble methods used by Govindarajan et al. (2024). This demonstrates that combining data balancing with optimal feature selection and deep learning enables highly accurate intrusion detection.

V. CONCLUSION

This study proposed a hybrid intrusion detection framework that integrates MI for feature selection with SVM and BPNN classifiers. Using the CICIDS2017 dataset, the system was further enhanced with SMOTE to address class imbalance, reducing the likelihood of biased classification. Six model configurations were evaluated using five key performance metrics: accuracy, precision, recall, F1-score, and ROC-AUC. Among them, the SMOTE + MI + BPNN model achieved the highest performance, significantly outperforming traditional IDS configurations and related studies. These results confirm that combining dimensionality reduction and class rebalancing techniques with deep learning can dramatically improve IDS effectiveness in detecting both known and unknown threats. These methods have greatly enhanced IDS performance in detection accuracy, scalability, and adaptability with minimal computation overhead. Lastly, the study achieved state-of-the-art performance using a hybrid deep learning model, offering a viable solution for real-world intrusion detection systems. Future work needs to emphasize on enhancing IDS models to detect attacks in real time and make them more resilient to resist new attacks, such as zero-day attacks and adversarial attacks. Class imbalance in IDS data sets is a reality because it has a direct impact on the detection accuracy as well as decreases false positives. Moreover, designing more energy-aware and lightweight IDS models specifically tailored for resource-limited setups, such as IoT and edge computing, will also play a vital role in making IDS more practical.

REFERENCES

[1] Ahmed, R., Zhang, Y., & Kim, J. (2023). Improving anomaly-based intrusion detection systems using hybrid SVM classifiers. *Journal of Information Security, 14*(2), 112–124.

[2] Alsaffar, A. M., Nouri-Baygi, M., & Zolbanin, H. M. (2024). Shielding networks: Enhancing intrusion detection with hybrid feature selection and stack ensemble learning. *Expert Systems with Applications, 239*, 121981.

[3] Afolabi, A., Ibrahim, M., Yusuf, R., & Adekunle, A. (2024). Network intrusion detection using knapsack optimization, mutual information gain, and machine learning (KOMIG-IDS). *Journal of Cybersecurity and Intelligence Systems, 6*(2), 85–101.

[4] Govindarajan, M., Arunachalam, R., & Singh, A. (2024). Performance analysis of mutual information-based intrusion detection using ensemble classifiers. *International Journal of Computer Applications, 182*(6), 10–17.

[5] Gupta, N., Sharma, A., & Choudhury, A. (2024). Cybersecurity threat detection using intelligent systems: A review of recent advances. *ACM Computing Surveys, 56*(3), 42–62.

[6] Hongchen, Y., Wei, Z., Chunying, K., & Yankun, X. (2025). A feature selection algorithm for intrusion detection system based on the enhanced heuristic optimizer. *IEEE Access, 13*, 33714–33729.

[7] Jain, A., & Kumar, V. (2023). Comparative study of IDS approaches: A shift from traditional to hybrid methods. *Information Systems Frontiers, 25*(1), 1–16.

[8] Liao, H., Tian, X., & Wang, R. (2023). Reducing false alarms in anomaly-based IDS using adaptive thresholding techniques. *IEEE Transactions on Information Forensics and Security, 18*, 270–282.

[9] Peng, H., Long, F., & Ding, C. (2023). Feature selection based on mutual information criteria: A new method. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(7), 1226–1238.

[10] Sarhan, M., Layeghy, S., & Portmann, M. (2021). Feature analysis for ML-based IoT intrusion detection. *IEEE Internet of Things Journal, 8*(5), 3452–3463.

[11] Yu, H., Zhang, W., Zhang, W., Kang, C., & Xue, Y. (2025). A feature selection algorithm for intrusion detection system based on the enhanced heuristic optimizer. Expert Systems with Applications, 265, 125860.

[12] Zhang, Y., Chen, X., & Li, W. (2024). A mutual information-enhanced ensemble framework for industrial control system anomaly detection. *Computers in Industry, 148*, 103832.

https://doi.org/10.5281/zenodo.16744756