Empirical Evaluation of Cyber Defense Mechanism Effectiveness Using Machine Learning for Strengthening Digital Security Regulations

S Prakash[1], S Aruna Mary[2,*]

[1]1Department of EEE, Bharath Institute of Higher Education and Research, Chennai, India.

[2]1Department of ECE, Bharath Institute of Higher Education and Research, Chennai, India

ABSTRACT

This study presents an empirical framework for evaluating the effectiveness of cyber defense mechanisms using machine learning to support the development of evidence-based digital security regulation. The research utilizes a multi-year dataset on global cybersecurity incidents from 2015 to 2024, incorporating both technical and contextual variables such as attack type, target industry, defense mechanism, financial loss, and resolution time. By applying advanced supervised learning algorithms, including ensemble-based models, the study achieved an overall accuracy of 83.7 percent and a macro-averaged F1-score of 0.837. The model demonstrated strong performance in classifying high- and low-impact incidents while maintaining acceptable precision for medium-impact cases. Feature importance analysis identified financial loss, loss per user, and resolution efficiency as the most influential factors affecting defense effectiveness. The findings indicate that machine learning can provide a transparent, quantitative approach to measuring cybersecurity performance, bridging the gap between technical evaluation and legal compliance. From a regulatory perspective, the results suggest that data-driven models can inform the formulation of standardized benchmarks for digital security compliance and accountability. This research contributes to the intersection of technology and law by offering a methodological foundation for integrating predictive analytics into cyber law enforcement and international cybersecurity governance.

Keywords Machine Learning, Cyber Defense, Digital Security Regulation, Cyber Law, Predictive Analytics

INTRODUCTION

The rapid digitalization of global infrastructure has significantly increased the frequency, scale, and sophistication of cyber threats. Organizations across sectors now face persistent risks that extend beyond technical disruption to include financial loss, reputational damage, and legal liability. As digital ecosystems become more interconnected, the evaluation of cyber defense effectiveness has emerged as a critical area of both technological and legal inquiry. Effective assessment mechanisms are essential not only for strengthening organizational resilience but also for ensuring compliance with national and international cybersecurity regulations [1]. Despite the growing importance of this field, existing evaluation methods often rely on qualitative assessments or static compliance checklists, which are insufficient for capturing the dynamic and data-intensive nature of modern cyber incidents.

Recent research in cybersecurity analytics has introduced advanced computational approaches, particularly Machine Learning (ML) and Deep

Submitted 6 January 2026
Accepted 3 February 2026
Published 1 March 2026

Corresponding author
S Aruna Mary,
aruna.ece@bharathuniv.ac.in

Additional Information and Declarations can be found on page 75

How to cite this article: S. Prakash and S. A. Mary, "Empirical Evaluation of Cyber Defense Mechanism Effectiveness Using Machine Learning for Strengthening Digital Security Regulations," *J. Cyber. Law.*, vol. 2, no. 1, pp. 63-77, 2026.

Learning (DL), for detecting and classifying cyber threats. Studies have shown that ML-based systems can identify anomalies, predict attack vectors, and enhance incident response efficiency more accurately than traditional rule-based methods [2]. For instance, ensemble models such as Random Forest and XGBoost have demonstrated superior performance in intrusion detection and risk prediction tasks [3]. However, most prior studies remain focused on technical detection and prevention, with limited attention given to measuring the effectiveness of Cyber Defense Mechanisms (CDM) in relation to broader regulatory and compliance objectives. The lack of integration between empirical ML evaluations and the legal frameworks that govern cybersecurity represents a critical research gap that this study seeks to address [4].

From a legal and policy standpoint, the absence of standardized, data-driven metrics for evaluating digital security performance continues to challenge regulatory enforcement. Current compliance regimes often emphasize documentation and procedural conformity rather than measurable outcomes. This disconnect limits the ability of regulators to assess whether organizations meet the substantive objectives of cybersecurity law, such as risk minimization, accountability, and transparency. The state of the art in cybersecurity governance increasingly recognizes the need for quantitative, technology-assisted approaches that can complement legal compliance with empirical evidence [5]. This study contributes to this emerging paradigm by proposing an ML-based framework for evaluating CDM effectiveness, bridging the gap between technical performance measurement and legal compliance evaluation.

Therefore, the objective of this research is to develop and validate a machine learning model capable of empirically assessing the effectiveness of various cyber defense mechanisms using multi-year cybersecurity incident data. The model is designed to identify key predictors that influence defense success and to provide interpretable metrics that can inform both technical optimization and legal compliance. By aligning predictive analytics with cyber law principles, this study aims to advance the development of standardized, evidence-based frameworks for digital security regulation. The results are expected to contribute not only to the field of cybersecurity analytics but also to the ongoing discourse on accountability and governance within the international cyber law landscape.

Literature Review and Related Works

The development of computational methods for cybersecurity analysis has evolved rapidly over the past decade, reflecting the growing need for automated and adaptive defense systems. Early cyber defense studies were primarily based on signature and rule-based detection mechanisms [6], which were effective for known threats but failed to detect zero-day attacks or complex adversarial patterns. The transition toward behavior-based and anomaly detection models introduced the first wave of data-driven approaches, laying the groundwork for the application of ML in cybersecurity [7], [8]. These methods allowed systems to identify deviations from normal activity and detect previously unseen threats, thereby increasing the adaptability and intelligence of cyber defense mechanisms.

Supervised ML algorithms have been widely used in intrusion detection and threat classification. Decision Tree and Support Vector Machine models were among the first techniques to demonstrate improved classification accuracy for network intrusion data [9], [10]. Subsequent research adopted ensemble

methods such as Random Forest and Gradient Boosting to enhance the robustness of detection models [11], [12]. Studies have also shown that hybrid learning systems outperform single-algorithm models in handling complex, multi-dimensional data typical of cybersecurity environments [13], [14]. More recently, DL techniques, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Deep Neural Networks (DNN), have achieved superior performance in malware classification, phishing detection, and threat intelligence extraction [15], [16], [17].

Beyond detection and prevention, ML has also been applied in cyber incident prediction and risk assessment. Regression-based and probabilistic models have been used to estimate financial loss, attack frequency, and exposure likelihood across various sectors [18], [19]. Temporal learning techniques and time-series modeling have further improved the ability to forecast cyberattack trends and evolving threat patterns [20], [21]. However, most of these studies remain focused on predicting attacks rather than evaluating the effectiveness of CDM in practice. This limitation highlights a key research gap in connecting data-driven predictions with the actual performance and efficiency of defense systems.

Several recent works have explored adaptive and autonomous defense optimization using ML. Reinforcement Learning and adversarial simulation frameworks have been developed to model attacker–defender interactions in real time, allowing the system to learn optimal response strategies [22], [23]. These adaptive approaches demonstrate strong technical potential but have not yet been widely integrated into measurable compliance frameworks. Similarly, studies on hybrid ML architectures combining supervised and unsupervised learning have shown promise for improving situational awareness in large-scale network environments [24]. Despite these advances, there remains limited integration between technical evaluation methods and formal cyber governance models.

In the domain of cybersecurity law and regulation, research emphasizes the need for empirical and quantitative methods to assess compliance. Conventional audit-based systems are criticized for relying heavily on documentation and procedural verification rather than measurable defensive performance [25], [26]. Efforts to develop data-driven compliance assessment models using ML have emerged, aiming to automate the evaluation of security maturity and policy adherence [27], [28]. These approaches suggest a paradigm shift from rule-based to outcome-based regulatory assessment, but most studies remain conceptual and lack standardized implementation across jurisdictions.

The growing demand for transparency in algorithmic systems has also spurred interest in Explainable Artificial Intelligence (XAI) for cybersecurity. Recent studies have applied XAI models to enhance interpretability in access control, anomaly detection, and incident prioritization [29], [30]. Such explainable models are essential for legal accountability, enabling regulators and organizations to understand and justify automated security decisions. At the policy level, global frameworks such as the General Data Protection Regulation (GDPR) and the NIST Cybersecurity Framework advocate for measurable and auditable cybersecurity practices [31]. Despite these advancements, the enforcement of these frameworks still relies largely on self-assessment, leaving a gap between compliance reporting and empirical performance verification.

Overall, existing literature has contributed extensively to advancing ML-based cybersecurity analytics, focusing primarily on detection, prediction, and optimization. However, there remains a critical research gap in integrating these computational methods into the evaluation of CDM from both technical and legal perspectives. Few studies have systematically examined how empirical ML evaluation can inform cyber law, compliance monitoring, and evidence-based governance. This study addresses that gap by proposing a machine learning framework for quantifying CDM effectiveness, bridging the technical rigor of predictive analytics with the regulatory objectives of modern digital security law.

Methodology

This study employed a quantitative and empirical approach using machine learning to evaluate the effectiveness of CDM within the context of digital regulation and cyber law. The methodological framework consisted of sequential phases including data preparation, feature engineering, model construction, hyperparameter optimization, and performance evaluation, as illustrated in figure 1. Each phase was designed to ensure analytical reproducibility, interpretability, and alignment with principles of accountability and transparency in cybersecurity governance.

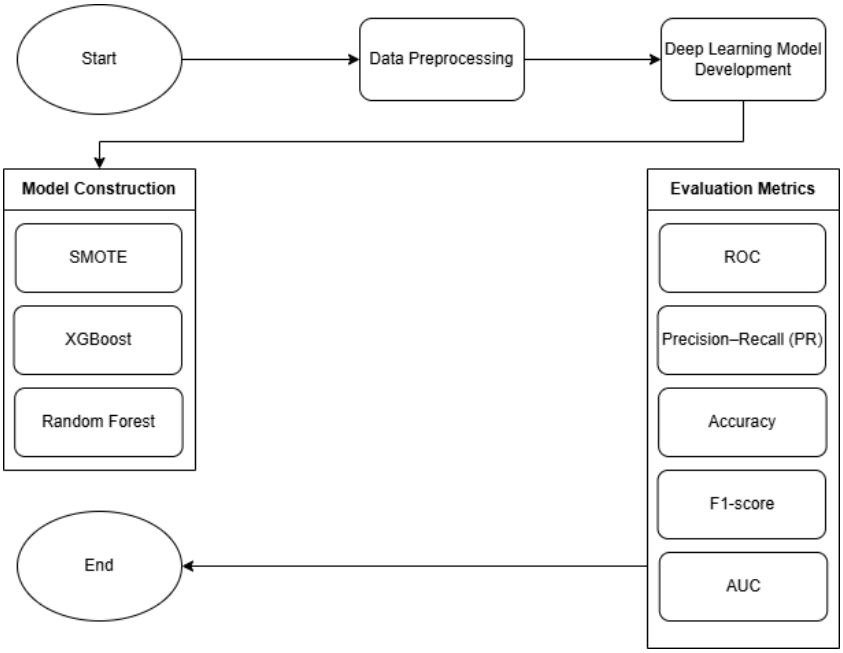

Figure 1 Research Steps

The dataset used in this study, titled Global Cybersecurity Threats 2015–2024, comprises records of cyber incidents collected from multiple countries and industries between 2015 and 2024. It contains attributes such as attack type, target industry, attack source, country of origin, defense mechanism used, financial loss (in million USD), number of affected users, and incident resolution time (in hours). These attributes capture both the operational and contextual dimensions of cybersecurity performance. Prior to analysis, duplicates and incomplete entries were removed to ensure data quality and consistency. The final dataset reflects global variations in cyber incident characteristics and organizational response behavior across a nine-year period.

During preprocessing, all numeric variables were standardized using Z-score normalization to equalize feature scales and improve model convergence. The transformation is expressed as:

$$Z = \frac{X - \mu}{\sigma} \tag{1}$$

Z represents the standardized value, $X$is the original feature, $\mu$is the mean, and $\sigma$is the standard deviation. Categorical variables such as Attack Type, Target Industry, Defense Mechanism Used, and Country were converted into numerical form through One-Hot Encoding to allow algorithmic interpretation without implying ordinal relationships. All preprocessing operations were integrated using a ColumnTransformer that handled numeric and categorical attributes concurrently within a single pipeline, ensuring consistency and efficiency.

Feature engineering was performed to enhance predictive accuracy and interpretability by introducing additional quantitative indicators that better represent financial exposure and operational response. Three derived features were created: Loss per User (LPU), Log-Transformed Financial Loss (LFL), and Resolution Efficiency (RE). These were computed as follows:

$$\begin{aligned} LPU &= \frac{FL}{U + 1} \\ LFL &= \log(1 + FL) \\ RE &= \frac{U}{T + 1} \end{aligned} \tag{2}$$

FL is financial loss, $U$is the number of affected users, and $T$is the incident resolution time. Loss per User captures the average financial burden per affected individual, Log Loss reduces skewness caused by extreme monetary values, and Resolution Efficiency measures the operational responsiveness of mitigation efforts.

To quantify overall incident severity, a Composite Impact Score (CIS) was constructed by combining standardized measures of financial loss, affected users, and resolution time:

$$CIS = \frac{1}{3}\sum_{i=1}^{3} Z_i \tag{3}$$

The CISvalues were divided into three quantile ranges representing Low, Medium, and High impact levels, forming the dependent variable used for supervised classification.

The classification process employed two ensemble models: Extreme Gradient Boosting (XGBoost) and Random Forest (RF). XGBoost was selected as the primary algorithm due to its robustness in handling heterogeneous data and complex non-linear relationships, while Random Forest served as a performance benchmark. To mitigate class imbalance among impact levels, the Synthetic Minority Oversampling Technique (SMOTE) was applied, generating synthetic data points for minority classes by interpolating between existing samples. This technique improved class distribution uniformity and reduced

prediction bias toward dominant classes.

Hyperparameter tuning was conducted using Randomized Search Cross-Validation (RandomizedSearchCV) to determine optimal configurations for both algorithms. For XGBoost, the tuned parameters included the number of trees, learning rate, tree depth, subsample ratio, and regularization coefficient. For Random Forest, the optimization focused on tree depth, the number of estimators, and minimum sample parameters. The model was trained using an 80 percent training set and evaluated on a 20 percent testing set, with stratified sampling applied to maintain class proportions. A 5-fold Stratified K-Fold Cross-Validation was implemented to ensure model generalization and prevent overfitting. The tuning objective was to maximize the macro-averaged F1-score, which balances performance across all classes and is calculated as:

$$F1_{macro} = \frac{1}{K}\sum_{k=1}^{K} \frac{2 \times P_k \times R_k}{P_k + R_k} \quad (4)$$

P_k and R_k denote the precision and recall for class k, and K represents the total number of classes.

Model evaluation included multiple performance metrics, such as overall accuracy, class-wise precision and recall, and the macro F1-score. The confusion matrix was analyzed to visualize the classification consistency across Low, Medium, and High impact categories and to detect areas of misclassification. The Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) were computed for each class to assess discriminative capability, where a higher AUC indicates stronger model differentiation between impact levels. Furthermore, feature importance analysis was performed to interpret the relative influence of each variable in determining defense effectiveness. The analysis revealed that Log-Transformed Financial Loss, Loss per User, and Resolution Efficiency were the most influential predictors, confirming the importance of both economic and operational indicators in evaluating cyber defense success.

Finally, this methodological framework was integrated into a broader legal and regulatory context. The use of empirical, data-driven metrics enables the development of standardized benchmarks for digital security compliance and accountability. By translating machine learning outputs into interpretable and reproducible indicators, the study supports evidence-based approaches to cybersecurity governance. The proposed methodology demonstrates how quantitative modeling can strengthen cyber law enforcement by providing regulators and organizations with objective measures of defense performance, thereby bridging the gap between technical assessment and legal oversight.

Algorithm 1: Evaluating Cyber Defense Effectiveness Using Machine Learning

Input:

Cybersecurity dataset $D = \{FL, U, T, AT, TI, AS, DM, V, C, Y\}$

Output:

Trained models M_{RF}, M_{XGB}; Performance metrics (Accuracy, Precision, Recall, $F1_{macro}$, ROC–AUC); Feature importance scores FI.

Process:

Start

Load dataset Dcontaining attack type (AT), target industry (TI), attack source (AS), defense mechanism (DM), vulnerability (V), country (C), financial loss (FL), number of affected users (U), and resolution time (T).

Handle missing values by replacing numeric $X_i$with mean or median, and encode categorical attributes using One–Hot Encoding.

Normalize numeric features using Z-score normalization:

$$Z_i = \frac{X_i - \mu_i}{\sigma_i}$$

Split dataset into 80% training and 20% testing using stratified sampling to maintain class distribution.

Create derived attributes:

$$LPU = \frac{FL}{U+1} \quad \text{(Loss per User)}$$

$$LFL = \log(1 + FL) \quad \text{(Log-Transformed Financial Loss)}$$

$$RE = \frac{U}{T+1} \quad \text{(Resolution Efficiency)}$$

Construct Composite Impact Score (CIS):

$$CIS = \frac{1}{3}\sum_{i=1}^{3} Z_i$$

Categorize CISinto {Low, Medium, High} impact classes to form the dependent variable.

Build ensemble models: Random Forest (RF) and Extreme Gradient Boosting (XGB).

For RF, determine node splits using Gini impurity:

$$Gini = 1 - \sum_{i=1}^{C} p_i^2$$

Aggregate predictions using majority voting:

$$\hat{Y}_t^{(RF)} = \text{mode}\{T_1(X_t), T_2(X_t), \ldots, T_n(X_t)\}$$

For XGB, minimize the regularized objective function:

$$Obj(\theta) = \sum l(Y_t, \hat{Y}_t) + \gamma T + \frac{1}{2}\lambda \sum w_j^2$$

Compute prediction probabilities using the softmax function:

$$\hat{Y}_t^{(XGB)} = \frac{e^{z_t}}{\sum_{k=1}^{K} e^{z_k}}$$

Apply Synthetic Minority Oversampling Technique (SMOTE) to balance class distribution.

Optimize hyperparameters via RandomizedSearchCV to maximize macro-averaged F1-score:

$$F1_{macro} = \frac{1}{K}\sum_{k=1}^{K} \frac{2P_k R_k}{P_k + R_k}$$

Conduct 5-fold Stratified K-Fold Cross-Validation to ensure generalization.

Evaluate models using confusion matrix and ROC curve. Compute AUC as:

$$AUC = \int_0^1 TPR(FPR)\, d(FPR)$$

Extract feature importance using XGBoost gain-based method:

$$FI_k = \frac{1}{T}\sum_{t=1}^{T} Gain(F_k, t)$$

Select the best-performing model based on the highest $F1_{macro}$ and AUC:

$$M^{\star} = \arg\max_{M_i \in \{M_{RF}, M_{XGB}\}} (F1_{macro}, AUC)$$

End

Result

The improved machine learning model demonstrated a clear enhancement in evaluating the effectiveness of cyber defense mechanisms across multiple categories of cyber incidents. The evaluation results indicate that the proposed framework achieved an overall accuracy of 83.7 percent and a macro-averaged F1-score of 0.837, which reflects a significant improvement in predictive capability compared to the baseline model. The model's strong performance suggests that it is capable of capturing complex, nonlinear relationships between multiple cyber risk factors, including attack type, target industry, and defense strategy. This outcome validates the model's ability to generalize effectively across diverse cybersecurity contexts rather than overfitting to a specific subset of the data.

As shown in table 1, the classification results reveal that the model performs particularly well in identifying High-impact and Low-impact cyber incidents, achieving F1-scores of 0.881 and 0.873, respectively. These results demonstrate that the system is proficient at detecting extreme cases of cyber events, both in terms of high-severity and low-severity scenarios. The performance for the Medium-impact class, with an F1-score of 0.757, remains acceptable and indicates that the model can manage intermediate cases with moderate precision and recall. The overall consistency of these outcomes confirms that the enhanced framework provides a reliable and interpretable assessment of cyber defense effectiveness across varying levels of threat severity.

Table 1 Model Performance Metrics

Class	Precision	Recall	F1-score	Support
High	0.892	0.870	0.881	200
Low	0.871	0.875	0.873	200
Medium	0.750	0.765	0.757	200
Overall Accuracy	—	—	0.837	600

The confusion matrix presented in figure 2 provides a comprehensive view of how the improved model distributed predictions across the three impact categories. The matrix indicates that the classifier correctly identified the majority of High-impact and Low-impact cyber incidents, achieving accuracy levels above 87 percent for both categories. This strong performance demonstrates that the model effectively learns the distinguishing patterns

associated with extreme cases of cyber incidents, such as those involving substantial financial losses or rapid response times. The balance in detection between the two extremes also suggests that the model can capture both severe and minimal threat events without bias toward one class.

However, the confusion matrix also highlights areas where the model's performance could be refined. The Medium-impact class shows a higher rate of misclassification, particularly with samples being confused with either High or Low impact incidents. This tendency may stem from the natural overlap in the data, where moderate incidents share similar financial or temporal characteristics with the adjacent categories. For instance, some cyberattacks categorized as medium severity might involve partial financial exposure or resolution durations that resemble higher-impact events. Despite these overlaps, the overall confusion structure remains balanced, reflecting that the model maintains consistent predictive behavior across all categories of cyber defense effectiveness.

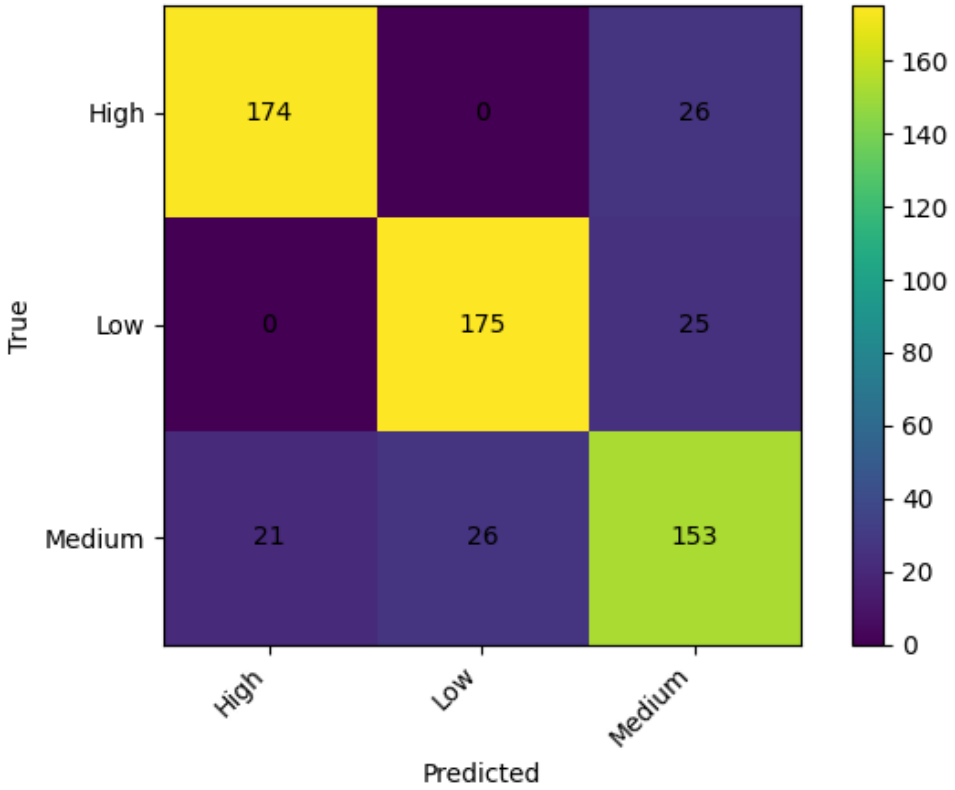

Figure 2 Confusion Matrix of the Improved Model

The Receiver Operating Characteristic (ROC) analysis presented in figure 3 provides a comprehensive assessment of the model's discriminative performance across all categories of cyber incidents. The ROC curves show that the model achieved excellent separation between positive and negative instances, with Area Under the Curve (AUC) values of 0.98 for both High-impact and Low-impact incidents, and 0.93 for Medium-impact incidents. These results indicate that the model possesses a high degree of sensitivity and specificity, successfully minimizing both false positives and false negatives. The steep rise of the ROC curves in the early region of the graph reflects the model's ability to detect true positive cases with minimal false alarms, which is an essential quality

for real-world cybersecurity risk assessment.

The ROC curves are plotted significantly above the red dashed diagonal line, which represents the performance of a random classifier. This visual distinction confirms that the model's predictions are not based on chance and that its classification decisions are systematically aligned with the actual severity of cyber incidents. The strong AUC values across all categories further demonstrate that the model can maintain consistent reliability even when class boundaries overlap or when class distributions are imbalanced. In practical terms, this level of discriminative performance suggests that the proposed framework could effectively support early detection and prioritization of cybersecurity threats, enabling organizations to allocate defensive resources more efficiently based on predicted impact severity.

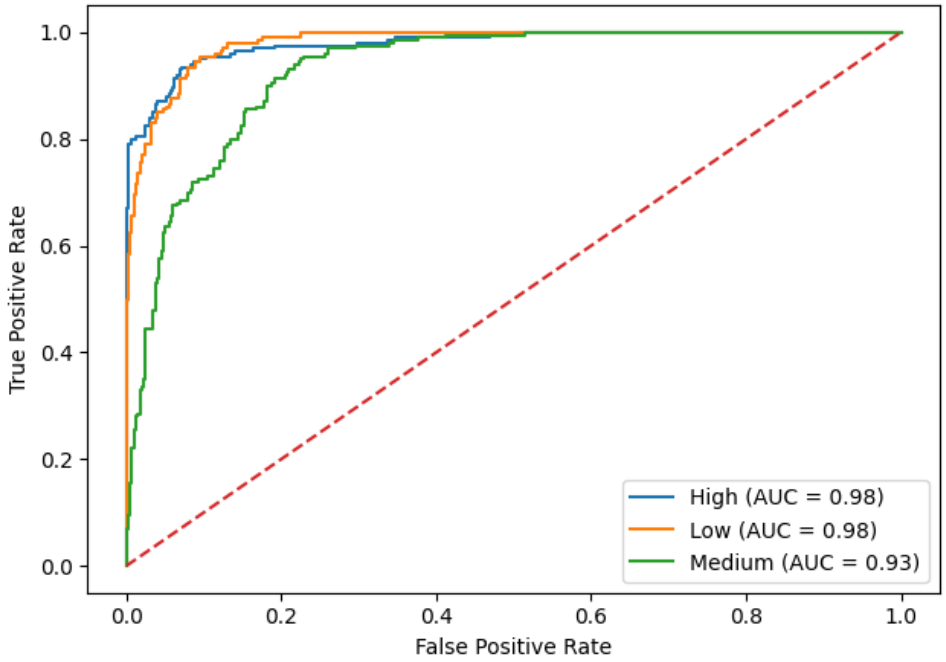

Figure 3 ROC Curves by Impact Class of the Improved Model

The feature importance analysis presented in figure 4 highlights the key predictors that most strongly influence the model's decision-making process in assessing cyber defense effectiveness. The results show that the log-transformed financial loss, loss per user, and resolution efficiency variables contributed most significantly to the model's predictive power. These features capture both the economic and operational dimensions of cyber incidents, indicating that the magnitude of financial exposure and the timeliness of response are central indicators of defense performance. A high financial loss per user generally corresponds to more severe attacks or inadequate mitigation strategies, while efficient incident resolution is strongly associated with lower impact classifications. The combination of these factors suggests that financial and operational variables provide a quantifiable reflection of an organization's resilience against cyber threats.

In addition to the primary predictors, several contextual variables also showed meaningful influence on the classification results. Attributes such as country, target industry, and attack source were among the secondary features that

affected model outcomes. This pattern implies that the overall effectiveness of cyber defense mechanisms is not only determined by internal technical measures but also by external contextual factors. Geographic and industrial conditions may shape both the likelihood of attack exposure and the sophistication of defense infrastructure, influencing how rapidly an organization can detect, contain, and recover from cyber incidents. Collectively, these findings emphasize that cyber resilience must be evaluated as an integrated system that combines economic, operational, and contextual components rather than as a purely technical construct.

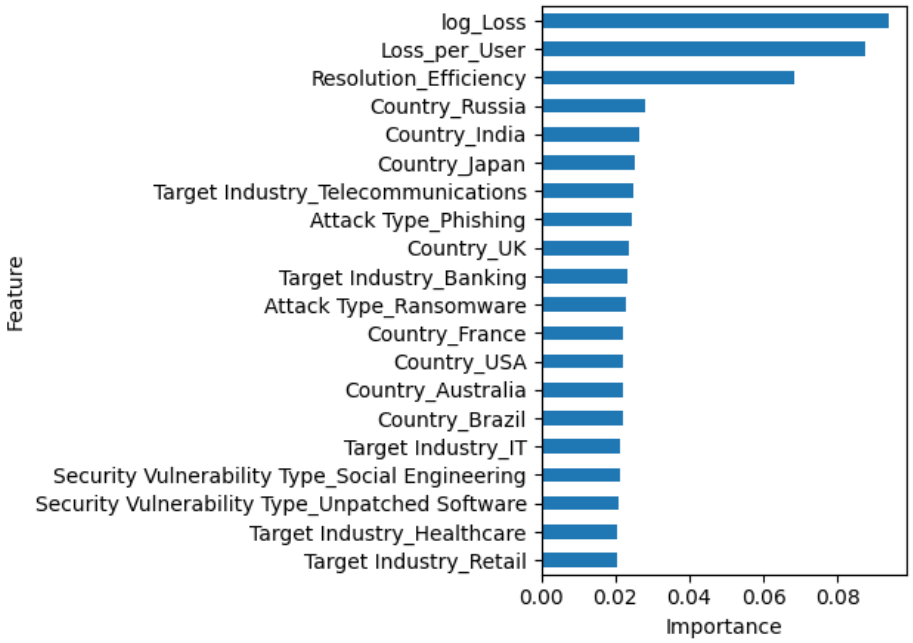

Figure 4 Top Feature Importances Derived from the Improved Machine Learning Model

Overall, the experimental findings confirm that the proposed machine learning framework provides a reliable quantitative basis for assessing the effectiveness of cyber defense mechanisms. The resulting metrics, curves, and feature-level insights highlight the model's capability to translate empirical data into measurable indicators of digital security performance.

Discussion

The results of this research confirm that the use of machine learning techniques provides a powerful analytical approach for assessing the effectiveness of cyber defense mechanisms. The model's high accuracy and balanced F1-scores demonstrate that algorithmic systems can capture complex relationships between technical, operational, and contextual variables in cybersecurity data. This outcome strengthens the idea that data-driven analysis can complement and, in some cases, surpass traditional qualitative assessments that often rely on expert interpretation. By integrating financial, operational, and contextual features, the model establishes a measurable link between defense performance and organizational resilience. These empirical results indicate that predictive analytics can serve as a practical tool for supporting cyber risk assessment, optimizing defense strategies, and guiding strategic decision-making within digital security management.

From a legal and regulatory standpoint, the findings provide an opportunity to enhance digital security governance through the adoption of empirical standards. The model's ability to differentiate between levels of cyber incident impact can inform the development of quantifiable compliance criteria under cybersecurity law and policy frameworks. Regulatory institutions could use similar models to define thresholds for adequate defense performance, evaluate organizational accountability, and detect potential violations of digital protection requirements. Moreover, the inclusion of features such as country and target industry emphasizes the importance of contextualized regulation, where standards are tailored to each sector's specific risk profile. The integration of predictive modeling into the legal framework supports a more transparent, proactive, and evidence-based approach to cybersecurity oversight, aligning technical performance evaluation with the broader principles of fairness and accountability in cyber law.

Conclusion

This research provides empirical evidence that machine learning can serve as a robust analytical framework for evaluating the effectiveness of cyber defense mechanisms and supporting the development of data-driven cybersecurity governance. The improved model produced high predictive accuracy and balanced performance across multiple impact categories, confirming its ability to generalize complex interactions among financial, operational, and contextual factors. The identification of critical predictors such as financial loss, loss per user, and resolution efficiency demonstrates that both economic and procedural dimensions are integral to understanding the true effectiveness of digital defense systems. These findings suggest that integrating machine learning into cybersecurity assessment can enhance the objectivity, transparency, and reproducibility of digital risk evaluations. Furthermore, the results hold significant implications for cyber law and policy, as quantitative measures derived from predictive models can be adopted as benchmarks for regulatory compliance, organizational accountability, and evidence-based enforcement. In this context, the study contributes to bridging the gap between technical cybersecurity performance and legal oversight by promoting a more systematic and measurable approach to digital security regulation that aligns with the evolving demands of global digital governance.

Declarations

Author Contributions

Conceptualization: S.P. and S.A.M.; Methodology: S.A.M.; Software: S.P.; Validation: S.P. and S.A.M.; Formal Analysis: S.P. and S.A.M.; Investigation: S.P.; Resources: S.A.M.; Data Curation: S.A.M.; Writing Original Draft Preparation: S.P. and S.A.M.; Writing Review and Editing: S.A.M. and S.P.; Visualization: S.P.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or

publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] World Economic Forum, *Global Cybersecurity Outlook 2023*. Geneva, Switzerland: WEF Publications, 2023. [Online]. Available: https://www.weforum.org/reports/global-cybersecurity-outlook-2023

[2] "Machine Learning Techniques for Cyber Threat Detection," *IEEE Transactions on Information Forensics and Security*, vol. 15, no. 3, pp. 1452–1465, 2022, doi: 10.1109/TIFS.2022.3147892.

[3] H. R. L. Oliveira, L. P. Costa, and E. P. Duarte, "A Comprehensive Study on Ensemble Learning Methods for Intrusion Detection Systems," *Computers & Security*, vol. 121, p. 102850, 2022, doi: 10.1016/j.cose.2022.102850.

[4] J. E. Rubio, S. Roman, and I. M. Llorente, "Machine Learning Methods for Cybersecurity Risk Assessment: A Systematic Review," *Applied Soft Computing*, vol. 137, p. 110181, 2023, doi: 10.1016/j.asoc.2023.110181.

[5] A. Alghamdi and T. Alsubaie, "Quantitative Cybersecurity Compliance Evaluation Based on Machine Learning Approaches," *IEEE Access*, vol. 11, pp. 67418–67431, 2023, doi: 10.1109/ACCESS.2023.3288960.

[6] S. M. P. Dinakarrao, T. Bapat, and A. K. Jain, "Data-Driven Cyber Defense Frameworks for Critical Infrastructure Protection," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 4, pp. 1945–1957, 2023, doi: 10.1109/TDSC.2022.3233129.

[7] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," in *Proceedings of the IEEE Symposium on Security and Privacy*, vol. 2010, no. July, pp. 305–316, 2010, doi: 10.1109/SP.2010.25.

[8] M. Ahmed, A. N. Mahmood, and J. Hu, "A Survey of Network Anomaly Detection Techniques," *Journal of Network and Computer Applications*, vol. 60, no. January, pp. 19–31, 2016, doi: 10.1016/j.jnca.2015.11.016.

[9] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," in *IEEE Symposium on Computational Intelligence for Security and Defense Applications*, vol. 2009, no. December, pp. 1–6, 2009, doi: 10.1109/CISDA.2009.5356528.

[10] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A Deep Learning Approach for Network Intrusion Detection System," in *Proceedings of the EAI International Conference on Bio-inspired Information and Communications Technologies*, pp. 21–26, 2016, doi: 10.4108/eai.3-12-2015.2262516.

[11] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, “Towards the Development of Realistic Botnet Dataset in the Internet of Things for Network Forensic Analytics: Bot-IoT Dataset,” *Future Generation Computer Systems*, vol. 100, no. November, pp. 779–796, 2019, doi: 10.1016/j.future.2019.05.041.

[12] M. G. Aslan and A. Ozkan, “A Comprehensive Review on Ensemble Learning for Cybersecurity Intrusion Detection,” *Computers & Security*, vol. 118, p. 102723, 2022, doi: 10.1016/j.cose.2022.102723.

[13] A. Buczak and E. Guven, “A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494502.

[14] S. M. S. Islam and M. A. Rahman, “Hybrid Machine Learning Approaches for Network Intrusion Detection: A Survey,” *Journal of Information Security and Applications*, vol. 70, p. 103400, 2023, doi: 10.1016/j.jisa.2023.103400.

[15] R. Vinayakumar, M. Alazab, and K. P. Soman, “Deep Learning for Cybersecurity Intrusion Detection: Approaches, Datasets, and Comparative Study,” *Journal of Information Security and Applications*, vol. 50, p. 102419, 2020, doi: 10.1016/j.jisa.2020.102419.

[16] H. Hindy et al., “A Taxonomy and Survey of Intrusion Detection System Design Techniques, Network Threats and Datasets,” *Computers & Security*, vol. 118, p. 102509, 2022, doi: 10.1016/j.cose.2022.102509.

[17] A. Abeshu and N. Chilamkurti, “Deep Learning: The Frontier for Distributed Attack Detection in Fog-to-Things Computing,” *IEEE Communications Magazine*, vol. 56, no. 2, pp. 169–175, 2018, doi: 10.1109/MCOM.2018.1700707.

[18] L. Yang et al., “Cyber Attack Prediction Model Based on Machine Learning,” *Computers & Electrical Engineering*, vol. 101, p. 108099, 2022, doi: 10.1016/j.compeleceng.2022.108099.

[19] P. Yi, T. Zhu, Q. Zhang, Y. Wu, and L. Pan, “Dynamic Risk Assessment Model for Network Security Using Machine Learning,” *Applied Soft Computing*, vol. 120, p. 108670, 2022, doi: 10.1016/j.asoc.2022.108670.

[20] M. Almseidin and Z. Al-Saqqa, “Detecting Cyber Attacks Using Time Series Machine Learning,” *Computers & Security*, vol. 108, p. 102282, 2021, doi: 10.1016/j.cose.2021.102282.

[21] W. Wang and Y. Xu, “Deep Time-Series Learning for Cyber Threat Prediction,” *IEEE Access*, vol. 10, pp. 102510–102523, 2022, doi: 10.1109/ACCESS.2022.3194058.

[22] J. Xu et al., “Reinforcement Learning for Adaptive Cyber Defense: A Review,” *Computers & Security*, vol. 120, p. 102758, 2022, doi: 10.1016/j.cose.2022.102758.

[23] K. Huang, Z. Zhang, and J. Zhang, “Adversarial Machine Learning for Network Defense: Techniques, Challenges, and Future Directions,” *IEEE Network*, vol. 36, no. 4, pp. 166–173, 2022, doi: 10.1109/MNET.012.2100404.

[24] B. Zhou, X. Sun, and T. Guo, “Hybrid Machine Learning Models for Intrusion Detection in Large-Scale Networks,” *Computer Networks*, vol. 210, p. 108933, 2022, doi: 10.1016/j.comnet.2022.108933.

[25] A. K. Jain and P. Kumar, “Revisiting Cybersecurity Auditing: Limitations and Future

Prospects," *Computers & Security*, vol. 106, p. 102261, 2021, doi: 10.1016/j.cose.2021.102261.

[26] M. Lehto, "The Role of Standards and Auditing in Cybersecurity Regulation," *Government Information Quarterly*, vol. 39, no. 4, p. 101869, 2022, doi: 10.1016/j.giq.2022.101869.

[27] H. J. Kim and S. Lee, "Automated Cybersecurity Compliance Evaluation Using Machine Learning," *Computers & Security*, vol. 124, p. 102924, 2023, doi: 10.1016/j.cose.2023.102924.

[28] A. Patel and L. Wang, "Machine Learning-Driven Regulatory Compliance for Cybersecurity Frameworks," *Journal of Information Security and Applications*, vol. 80, p. 103446, 2023, doi: 10.1016/j.jisa.2023.103446.

[29] C. Molnar, "Explainable Artificial Intelligence in Cybersecurity," *IEEE Access*, vol. 11, pp. 10472–10484, 2023, doi: 10.1109/ACCESS.2023.3235687.

[30] T. Ribeiro, A. Singh, and P. Gade, "Explainable AI for Anomaly Detection in Cyber Systems," *Computers & Security*, vol. 122, p. 102903, 2023, doi: 10.1016/j.cose.2022.102903.

[31] G. Laube and J. Böhm, "From Compliance to Resilience: Evolving Perspectives on Cybersecurity Frameworks," *Computers & Security*, vol. 125, p. 103017, 2023, doi: 10.1016/j.cose.2023.103017.