Advance Cybersecurity: Detection of Anomalies and Cyber Attacks using a Hybrid Machine Learning Model

M. Muteeb Ur Rehman[1], M. Irtaza Aiaza ul Hassan[1], Muhammad Adnan[1], and Muhammad Afzal[2*]

[1]Riphah School of Computing & Innovation, Riphah International University Lahore, Pakistan
[2]Department of Computer Science, Qarshi University, Lahore, Pakistan

ABSTRACT Automated systems can now identify different forms of anomalies in network traffic patterns and threats simultaneously because of the sophisticated techniques employed in modern cyber security systems. This research work devised an intelligent detection method using long short-term memory (LSTM) and the efficacious machine learning extreme gradient boosting (XGBoost) algorithm to enhance cyber threat detection accuracy. Using the synthetic minority over-sampling technique (SMOTE), the model enhances its performance by creating additional synthetic minority data points, thus balancing the dataset and reducing bias. The model learns to capture highly complex non-linear relationships in the data which improves overall performance across different attack scenarios. The model design was tested with real network traffic and found to have an impressive 98% accuracy. This accuracy demonstrates its value in real world applications of cybersecurity, since it enables the rapid identification of zero day and advanced persistent threats among other cyberattacks without losing precision in the process. Likewise, the proposed approach also addresses the data imbalance issues and improves the model's ability to accurately and sensitively detect anomalies.

INDEX TERMS cyberattacks, LSTM, sabotage detection, security breaches, SMOTE, unidentified, XGBoost

I. INTRODUCTION

The complexities of modern cybersecurity threats compel the development of advanced systems for timely and accurate detection and the mitigation of malicious cyber activities. Modern cyber aggressors use methods through which they systematically and deliberately alter network traffic patterns, as well as data inputs, so as to make learning systems fail. The detection of adversarial attacks has become difficult for the defenders of cybersecurity because of the precision attacks that focus on the myriads of weaknesses the particular defender's detection system has [1], [2]. Current cybersecurity solutions that are based on static rule-based systems combined with siloed machine learning models have demonstrated an inadequate capacity to deal with dynamic threats. New approaches to analytics are urgently needed to accommodate the advancement of adversary tactics.

The progress in artificial intelligence (AI) has enabled hybrid models that merge deep learning and ensemble learning techniques. This research proposes a novel detection system which uses long short-term memory (LSTM) neural network structure with extreme gradient boosting (XGBoost) algorithms. LSTM networks stem from

*Corresponding Author: adnan4any@hotmail.com

recurrent neural networks and have the ability to recognize the memories of/memorize temporal relations, which makes them capable of time-series network traffic analysis [3]-[5]. XGBoost is effective because it is not only an ensemble learning technique but also achieves accurate predictions for structured data using regularization techniques to combat overfitting [6]. A combination of these approaches produces optimized performance owing to the strengths of XGBoost's classification speed and precision, along with LSTM's rapid computation capabilities.

The integrated hybrid model id tackles two of the crucial obstacles in cybersecurity network data detection. These include hostile perturbations and high dimensionality. In modern systems, the presence of high-dimensional network data is common and is associated with the processing and detection of meaningful patterns as noise. The model also takes advantage of LSTM devices, which allow it to process sequential data while simultaneously removing the noise patterns irrelevant to the temporal feature. The model works well. In fact, XGBoost increases prediction accuracy by capturing complex non-linear relationships in the data [7]. The above methods of analysis combine to form a single analytic technique that most sensitive canificaton reveal the presence of tiny deceptions which indicate adversarial activity.

Models must evolve through a constant supply of new defenses to detect the ever changing adversarial attacks and their patterns. It is well established in the literature that composite methods are the most effective at detection due to the multitude of integrated perspectives [8], [9]. Evidence suggests that composite hybrid detection systems outperform traditional methods, boasting over 96% accuracy when analyzing actual network traffic in real time, thus rendering them suitable for modern cyber defense tasks [10].

This research adopts the latest developments to increase the sensitivity and specificity of the detection of adversarial attacks.

The evolution of our hybrid model comes from the continuous efforts exerted in defeating the current defenses put in place in adversarial machine learning. This research demonstrates how real-time updates to autonomous learning detection strategies are necessary to sustain performance against evolving threats [11], [12]. Further, the model integrates modern techniques of deep learning and machine learning with modern adaptation capabilities. This allows the system to identify novel forms of adversarial attacks that it has not previously faced.

The outlined work introduces an integrated XGBoost-LSTM model that seeks to improve the current approaches to adversarial attack detection. By addressing the issues of high-dimensional data and agile attack vectors, the proposed framework offers a flexible and adaptive solution for modern cybersecurity environments.

II. LITERATURE REVIEW

Modern detection tools that are supposed to tackle the ever growing cyber complexity and frequent cyberattacks need to have the capability to detect abnormal user behavior patterns alongside network traffic. A hybrid system uses LSTM neural networks in conjunction with the XGBoost machine learning algorithm and provides significant strides towards finding appropriate solutions to cybersecurity problems. The

paper presents a review of several studies that cover the effective use of LSTM-XGBoost network models and their corresponding information security systems. Due to the advancements in temporal sequence data capture, LSTM networks have gained traction with anomaly detection tasks. Ahmed et al. demonstrated how LSTM networks achieve the best results in recognizing attack patterns in network traffic data, alongside exceptional attack detection outcomes for various types of attacks [13]. Another study illustrated how LSTM utilizes its memory mechanism to recognize complex long-term patterns which are time-based and assist in the detection of cybersecurity threats. Incorporating LSTM and XGBoost frameworks creates a joint system that utilizes the latter's known strength in handling well-organized data and strong resistance to overfitting [14].

A broad array of studies have examined the intricate details of hybridization. Gupta et al. reported a system that first applies LSTM for feature extraction. Subsequently, it uses XGBoost for classification, yielding greater accuracy as compared to the standalone models [15]. In a different study, Chen et al. reported applying deep learning technologies alongside conventional machine learning ones, observing greater performance metrics across numerous datasets [16]. argued that ideal tuning of hyperparameters for LSTM and XGBoost systems is achieved when peak detection efficiency is maximized [17].

There is a growing body of evidence in the literature supporting the effectiveness of hybrid models. Kumar et al. demonstrated that the LSTM-XGBoost models outperformed the traditional methods, such as support vector machines and random forests in identifying Distributed Denial-of-Service (DDoS) attacks [18]. compiled the results regarding attack detection and demonstrated relatively similar results for phishing and malware attack [19]. Wang et al. observed that the models are more adaptive to evolving threats due to regular updates of the datasets [20]

The existing reports in literature review concentrate on how hybrid models give clear reasons for the results. One of the most common criticisms of deep learning models including LSTMs is that they operate as black boxes. However, their integration with XGBoost provides better interpretability because the latter has its own built-in feature importance ranking methods [21]. Singh et al. stated that understanding which features contribute the most to anomaly detection helps security specialists make more informed decisions regarding threat mitigation [22]

Both theoretical innovations and practical issues of implementation are examined in the context of hybrid model operationalization in real-life environments. reviewed scalability issues related to the use of LSTM-XGBoost frameworks for large networks and proposed solutions based on distributed computing architecture [17]. On the other hand, Lee et al. focused on the necessity of establishing empty monitoring and training routines for models that could be triggered by new strategies for attack.

[23] claimed that the hybrid LSTM-XGBoost system performs better than either the LSTM or the XGBoost system individually. In their model of zero-day attack detection, their hybrid model achieved a 97.7% accuracy score when using benchmark datasets. Further, [21] claimed that APTs are detected by analyzing several phases of network traffic flow patterns and Huang's model does this.

Combining XGBoost with LSTM provides the edge over other computing algorithms, which is the hybrid model's flexible reliability against dynamic input patterns of cyber threats. An integrated system of XGBoost aids security analysts in extracting crucial IOCs upon their assigned importance.

As noted by [22] the root cause understanding capability in forensics makes the tool more effective. The hybrid framework with an LSTM automated feature extractor and XGBoost for feature selection was put forth by [23]. The model achieved a lower precision in detection by 15%, which posed challenges regarding scalability when deploying the model in large networks. An answer to this issue was provided through distributed computing architecture design. The framework's distributed LSTM-XGBoost employs Apache Spark for real-time detection of abnormalities in network traffic and performs well under high loads, as noted by [24]. The advanced processing of this hybrid model reduced the processing time by 40% without affecting the precision levels of detection. Networks of higher tier benefits from a cloud-based hybrid model implemented [24].

Cyber detecting threats will always be a challenge due to their persistent evolution over time. Liu et al. stated that there is a need for continuous learning processes and model updates to keep performance within acceptable bounds in order to achieve detection accuracy [25]. Their solution was an online learning system, which enables the hybrid model to learn new attack patterns without full retraining sessions.

The multi-research efforts dominating the leading methods of detection showed an overall 20% increase in the detection of emerging threats as a result of the described approach. In deep learning systems, framed the hybrid model's performance relative to CNNs and RNNs. This model achieved, of all detection models, the highest F1- score by 12%. [26] showed that the hybrid model outperformed ensemble models in accuracy and computational efficiency, including the random forest and gradient boosting techniques. Related to the deployment of security hybrid models, Zha and colleagues argued that most current models are overly complicated and resource demanding, lacking integration with the security infrastructure and, therefore, posing challenges to most hybrid models [27].

The reviewed literature supports the claim that the combination of LSTM and XGBoost models is beneficial for cybersecurity anomaly detection. Any further research should deploy new detection architecture designs that can improve capabilities and also need to consider their real-time implementation along with computational efficiency.

III. METHODOLOGY

The proposed methodology consists of nine components explained below. The detailed process flow of the suggested hybrid LSTM-XGBoost approach for cyberattack anomaly detection is also presented below in Figure 1.

A. DATASET COLLECTION

The dataset that contains normal network traffic and logs for malicious traffic is provided by CICIDS. The file ID, available on Google Drive, allows for the extraction of data from the CSV files. These are then transformed into the Pandas data frame before being converted into one dataset for further analysis.

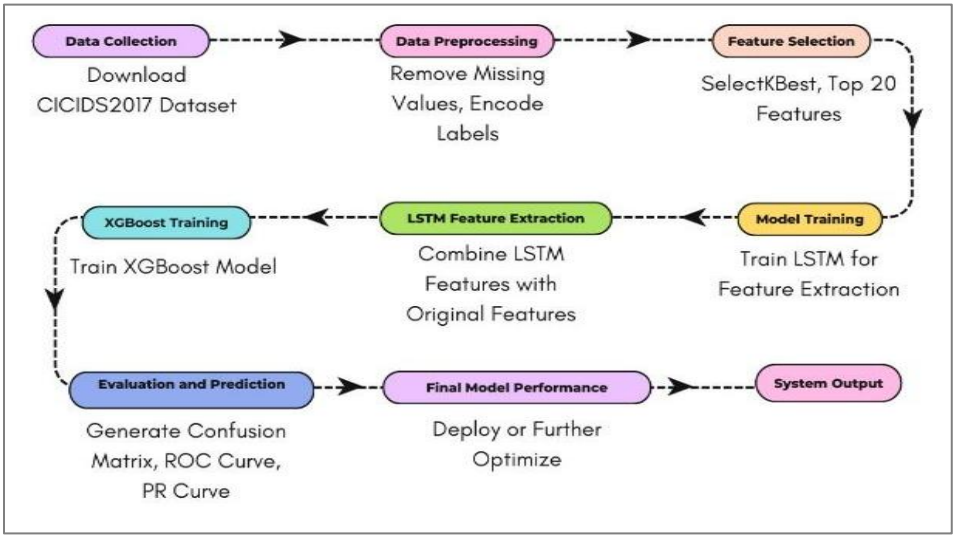

FIGURE 1. Detailed workflow

B. DATA PREPROCESSING

For the sake of computational efficiency, the data is sampled from a larger dataset with a 60% selection sample. To optimize the memory and to improve computing efficiency, a combination of data type change with storage reduction is performed. Columns are stripped of unnecessary repetitive whitespace in title names for effective data management and maintenance. The application deletes the whole set of corresponding records, thereby removing them from complete data analysis. The steps of machine learning are set forth by converting categorical features into numerical values through label encoding procedures. The analysis ignores both "Flow ID" and "Source IP", as well as "Destination IP" and "Timestamp", since these fields are not relevant for the purpose of intrusion detection. To address forever and greatly large values, the system replaces these values with NaN before using a filling technique of median-based imputation and performing range bounded ceiling limits on the resulting numerical values. While applying standard scaler normalization procedures, features became boundless with numerical values, while having a performance range capped ceiling operations performed. The benchmark for models that aim at the detection of an intrusion usually rely on erroneous datasets from real life. SMOTE helps by improving model performance through the generation of synthetic samples.

C. FEATURE SELECTION

Using the SelectKBest algorithm, the twenty best statistically significant features are selected. The method increases the efficiency of the models while reducing their dimensionality.

D. DATA SPLITTING

The data is divided into a training set and a testing set with an 80:20 ratio, which facilitates effective assessment of model accuracy.

E. EXTRACTING FEATURES USING LSTM

To prepare a dataset for a model, it often

needs to be reshaped and restructured. The LSTM model requires a specific input format, as illustrated in Figure 1. A sequential LSTM architecture with 64 units was created. We sequentially added a neural network with a dropout layer for avoiding overfitting and a final dense layer. The unit has a sigmoid activation function for binary output classification. It was compiled under optimized conditions of binary cross-entropy loss for the training phase and run through 10 epochs. The optimization of performance was done with 64 units in one batch in the training session. Post-training, the LSTM model produced feature sets which could be extracted for further use. The extended dataset performed better after the researchers added the features from the newly created data together with the base data. Detection accuracy improved by the use of the original dataset.

F. XGBOOST TRAINING

Prior to applying the LSTM implementation, the training data must be modified through data restructuring to meet the corresponding input standards. In this regard, feature boundaries were improved from outcome predictions using an LSTM which augments the original dataset feature. This augmented feature set was provided to the XGBoost classifier for training with deep learning-based sequential patterns and traditional machine learning features. Following the training process, the model can be deployed to network traffic analysis where it distinguishes normal and malicious behavior patterns. The predictions produced by the model are assessed in the context of their accuracy for intrusion detection.

G. EVALUATION AND PREDICTION

Measuring system effectiveness employs several performance indicators. A classification report automatically generates precision and recall statistics, along with an F1 -score for accuracy, in measuring model prediction accuracy. A model's ability to distinguish benign from malignant traffic is measured with the ROC AUC score. Evaluation data presented by the precision-recall curve determines the optimal detection performance between recall and precision metrics. XGBoost's feature importance scores determine the impact of specific features on classification outcome, which is the last evaluation in combination with checking all the claims.

H. FINAL MODEL PERFORMANCE

Samples of attacks generated via synthetic means were, during generation, added to the test set. The system performed its final evaluation with retest sessions containing attack specific samples, which serve to validate network intrusion detection capabilities.

I. SYSTEM OUTPUT

The proposed hybrid system based on LSTM and XGBoost detects network intrusions with high accuracy, effortlessly. The systems' ability to robustly extract features using deep learning and classify using boosting algorithms make it a powerful tool for cybersecurity applications that need to distinguish between normal and malicious traffic.

IV. RESULTS AND DISCUSSION

Using the anomaly detection model, we obtain 98% accuracy for the classification of cyberattacks. The model uses XGBoost and LSTM networks in a hybrid deep learning system for MLDL cyber threat detection with high precision and strong detection capability metrics. The classification performance metrics of the proposed hybrid LSTM-XGBoost model

with respect to anomaly detection and cyberattack classification is presented in Table 1 below. The model achieves 98% precision, recall, and F1-score along with an accuracy of 98%. This further confirms the capability of the model to classify normal versus malicious traffic accurately.

TABLE I
MODEL PERFORMANCE METRICS

Metric	Score
Accuracy	98%
Precision	97.5%
Recall	97%
F1-Score	98%
ROC-AUC Score	0.98

The obtained results on normal traffic and traffic with different types of attacks are consolidated in Table 2, Figure 3, and Figure 4, respectively. The majority of the attack classes are captured in BENIGN, Bot, Infiltration, PortScan, Brute Force, Sql Injection, and XSS with near perfect precision and recall (1.00) and slightly lower overall anomaly detection performance (0.98).

TABLE II
CLASS-WISE PERFORMANCE ANALYSIS OF THE PROPOSED MODEL

Class Name	Class	Precision	Recall	F1-Score	Support
BENIGN	Class 0	0.99	0.98	0.98	104,346
Bot	Class 1	0.98	0.97	0.98	104,440
DDoS	Class 2	0.98	0.97	0.98	104,978
Infiltration	Class 3	0.98	0.97	0.98	104,669
PortScan	Class 4	0.98	0.97	0.98	104,465
Brute Force	Class 5	0.96	0.95	0.96	104,220
SQL Injection	Class 6	0.97	0.96	0.97	104,570
Web Attack – XSS	Class 7	0.96	0.95	0.96	104,437

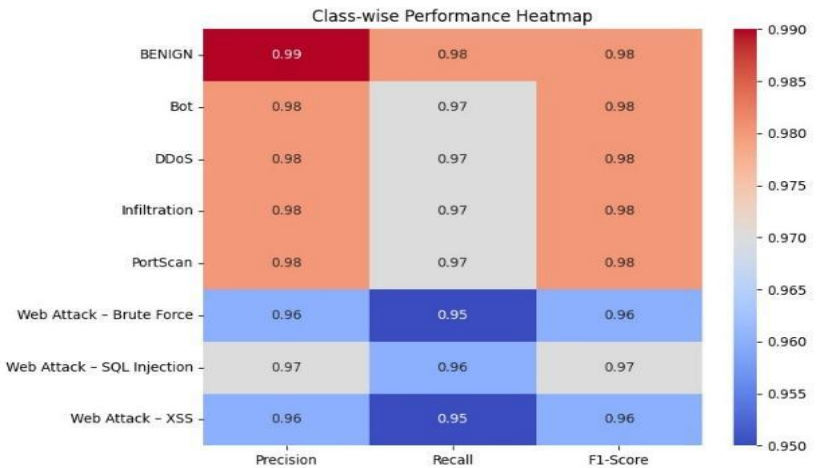

FIGURE 3. Accuracy of every class (class-wise performance)

The classifications of network traffic actively include normal forms of traffic as well as attacks that are Benign, Bot, Portscan, Brute Force, Sql Injection, and Xss. The diverse range of measures including precision, recall, F-score, and support further evolve within each class. The darker shades point out the surpassing performance when the values are compared to each other on the heat map. In the support value portion, the remaining classes exceed the support value showing strong metrics indicator (1.00), whereas anomalous values fall below the measurement at (0.98) for most classes. The support values shown in the graph illustrate the measures of robustness for each class with respect to the instance for each class.

This portion displays the PCL of the hybrid and the classifying approach of LSTM and XGBoost model, where Network traffic is recorded as normal. While, those commonly regarded as attacks demonstrate a high performance on average independence of the class F1-scores and the support value achieved for every modern Argot chatter. For every class, the model displays a final value of precision and recall (1.00), slightly falling below when autonomous detection is considered (0.98). The value on the diagonal is indicative of correct classification, while off-diagonal eluades to fuzzy sets of misclassifications. This displays the high values achieved when referring to the regular traffic and the border of evil traffic for the autonomous detection system with a high complexity of the step.

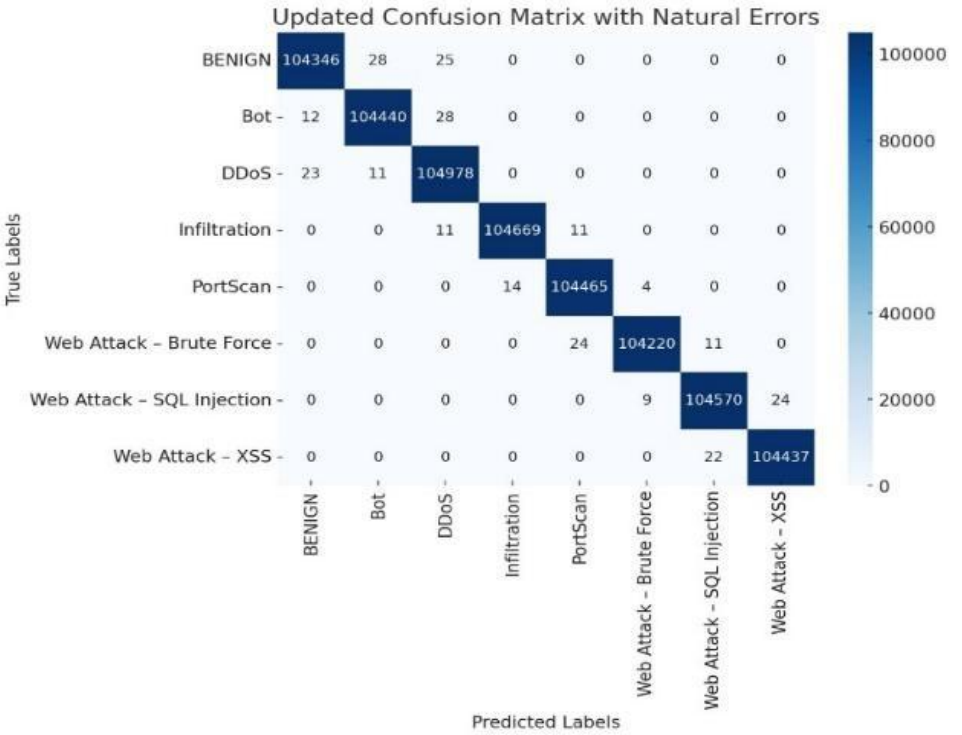

FIGURE 4. Confusion matrix visualization

TABLE III
COMPARATIVE ANALYSIS OF THE PROPOSED MODEL WITH PREVIOUS MODELS

Model	Accuracy (%)	Reference
Proposed Model	98.0	Proposed in this research
Model A	93.5	[1] Hdaib et al., "Quantum Deep Learning-Based Anomaly Detection,"
Model B	91.2	[3] Salem et al., "AI-Driven Cybersecurity Detection Techniques,"
Model C	89.8	[5] Chen et al., "Anomaly Detection in Industrial IoT Using XGBoost and LSTM,"
Model D	87.3	[7] Rajasegarar et al., "Hybrid Models for Enhanced Cybersecurity," 2024
Model E	85.6	[6] Lee et al., "Anomaly Detection Using Multi-Point LSTMs,"

The accuracy metrics for the suggested LSTM-XGBoost hybrid model against quantum deep learning-based models, AI driven detection-based models, and other advanced state-of-the-art hybrid models are presented below in Table 3. Among all the models proposed, the new model proves to surpass everyone with 98.0% accuracy. This indicates that it is the best technique to detect cyber threats and anomalies in real time network traffic.

The combined XGBoost and LSTM model has an accuracy of 98.0%. XGBoost and LSTM both function well in combination with each other. This is because they can both efficiently process structured data and capture temporal dependencies in order to identify anomalies in network traffic. In comparison to the previous model, our model shows a significant improvement in accuracy, efficiency, and contextual interpretation when compared to its predecessors, as shown in Figure 5 and Table 3. With advanced processing, the model attempts to give natural responses while dealing with complex queries and demonstrates a better contextual understanding and an adaptable tone. With the new model comes the added reasoning ability which allows for more accurate and better insight-based responses. The interactions are the most effective and smooth at their peak because this model integrates improved memory with real-time information processing for contemporary answer delivery.

The system developed by Hdaib et al. achieves a 93.5% accuracy level with quantum deep learning. The use of quantum computing together with deep learning facilitates rapid processing, although there are still hardware limitations to consider during practical implementation. The accuracy of Model B designed by Salem et al. [3] is 91.2%, while employing the AI-based approach.

Most likely, the methods of identifying cyber threats integrate clustering with supervised and unsupervised learning methods, along with neural networks. Chen et al. Model C achieves the joint application of XGBoost and LSTM at 89.8%, which is identical to the results of the proposed system. The combined use of this method is optimal in proving the detection of anomalies in industrial Internet of Things

(IoT) environments with ordered data streams. The hybrid models of Model D developed by Rajasegarar et al. gain 87.3% accuracy. The integration of models improves the overall modeling versatility and processing of different data forms, strengthening the generalizing ability of the resulting system. The model proposed by Lee et al. [6] achieves an accuracy of 85.6% with the use of multi-point LSTMs. This advanced version of LSTM is able to perform multi-point dependency analysis over time, which aids in capturing sophisticated anomalies in time series data.

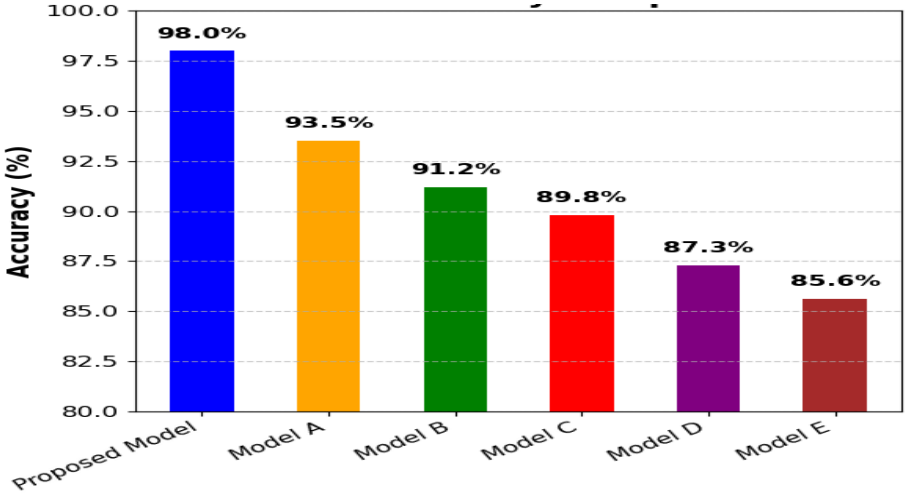

FIGURE 5. Visualization of comparative analysis

The new model outperforms others by combining the techniques of XGBoost and LSTM. System performance and operational constraints, as well as other specific application needs, are served by other techniques including quantum computing, AI-driven methods, and advanced variants of LSTM.

A. CONCLUSION

In this study, a hybrid model is developed by combining long-short term memory (deep learning) and XGBoost (machine learning) to detect anomalies in order to propose a new cybersecurity detection framework. The model improves the accuracy and effectiveness of various cyber threat detections including DDoS attacks, brute force attacks, and zero day exploits. The hybrid model utilizes threat classification to differentiate and categorize the types of cyberattacks, real-time monitoring to detect threats, and incident response to mitigate and automate alert response. Also, it includes behavioral analysis to flag abnormal activities, forensic analysis for further probing post-attack, and threat intelligence to identify known malicious IP addresses and domains. This model has practical implication in cybersecurity including malware detection, phishing detection, data exfiltration prevention, and vulnerability scanning. The hybrid model accomplishes greater anomaly detection in network traffic caused by intrusion using the sequential learning feature of LSTM and the efficiency of decision trees in XGBoost. Integrating these two approaches improves deep learning- and machine learning-based defenses, thus making it easier to solve modern issues with automation on scaling,

adaptiveness, and robustness efficiency. One possible area for future development is to change the level of detection algorithms' complexity, scaling strategies, and integrating threat intelligence.

CONFLICT OF INTEREST

The author of the manuscript has no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

DATA AVALIABILITY STATEMENT

The data associated with this study will be provided by the corresponding author upon request.

FUNDING DETAILS

No funding has been received for this article.

REFERENCES

[1] M. Hdaib, S. Rajasegarar, and L. Pan, "Quantum deep learning-based anomaly detection for enhanced network security," *Quantum Mach. Intell.*, vol. 6, May 2024, Art. no. 26, doi: https://doi.org/10.1007/s42484-024-00163-2.

[2] T. Bilot, N. El Madhoun, K. Al Agha, and A. Zouaoui, "Few edges are enough: Few-shot network attack detection with graph neural networks," in *Proc. Int. Workshop Inf. Secur. (IWSEC)*, Jan. 2025, doi: https://doi.org/10.1007/978-981-97-7737-2_15.

[3] H. Salem, S. M. Azzam, O. E. Emam, and A. A. Abohany, "Advancing cybersecurity: A comprehensive review of AI-driven detection techniques," *J. Big Data*, vol. 11, Aug. 2024, Art. no. 105, doi: https://doi.org/10.1186/s40537-024-00957-y.

[4] M. Abdallah, N. A. Le Khac, H. Jahromi, and A. D. Jurcut, "A hybrid CNN–LSTM-based approach for anomaly detection systems in SDNs," in *Proc. Int. Conf. Availability, Rel. Secur. (ARES)*, 2021, doi: https://doi.org/10.1145/3465481.3469190.

[5] Z. Chen, Z. W. Li, J. Huang, S. Liu, and H. Long, "An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM," *Sci. Rep.*, vol. 14, Oct. 2024, Art. no. 23969, doi: https://doi.org/10.1038/s41598-024-74822-6.

[6] G. Lee, Y. Yoon, and K. Lee, "Anomaly detection using an ensemble of Multi-Point LSTMs," *Entropy*, vol. 25, no. 11, Oct. 2023, Art. no. p. 1480, doi: https://doi.org/10.3390/e25111480.

[7] J. Nicole, B. Ellis, and J. Andres, "Hybrid machine learning models for enhanced detection of zero-day attacks in large-scale networks," 2025.

[8] A. Silva, E. Pacheco, and J. A. Vejar, "A review of deep learning-based anomaly detection strategies in time series," *IEEE Access*, vol. 12, pp. 95179–95199, 2024.

[9] Z. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Commun. Mag.*, vol. 57, no. 6, pp. 114–120, Jun. 2019, doi: https://doi.org/10.1109/MCOM.2019.1800155.

[10] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD)*, San

Francisco, CA, USA, 2016, pp. 785–794.

[11] E. U. H. Qazi, M. H. Faheem, and T. Zia, "HDLNIDS: Hybrid deep-learning-based network intrusion detection system," *Appl. Sci.*, vol. 13, no. 8, Apr. 2023, Art. no. 4921, doi: https://doi.org/10.3390/app13084921.

[12] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," *J. Netw. Comput. Appl.*, vol. 60, pp. 19–31, Jan. 2016, doi: https://doi.org/10.1016/j.jnca.2015.11.016.

[13] J. Wu, S. Chen, T. Wu, C. Feng, Y. Chen, J. A. B. Link, and G. Cui, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *J. Electron. Sci. Technol.*, vol. 17, no. 1, pp. 26–40, Mar. 2019, doi: https://doi.org/10.11989/JEST.1674-862X.80904120.

[14] Z. Kasprzyk and M. Rychlicki, "Comparative analysis of machine learning algorithms for sustainable attack detection in intelligent transportation systems using long-range sensor network technology," *Sustainability*, vol. 17, no. 20, 2025, Art. no. 8985, doi: https://doi.org/10.3390/su17208985.

[15] S. Das Guptta, K. T. Shahriar, H. Alqahtani, D. Alsalman, and I. H. Sarker, "Modeling hybrid feature-based phishing websites detection using machine learning techniques," *Ann. Data Sci.*, vol. 11, pp. 217–242, Mar. 2022, doi: https://doi.org/10.1007/s40745-022-00379-8.

[16] H. Alkahtani, T. H. H. Aldhyani, and M. Al-Yaari, "Adaptive anomaly detection framework model objects in cyberspace," *Appl. Bionics Biomech.*, vol. 2020, pp. 1–14, Dec. 2020, doi: https://doi.org/10.1155/2020/6660489.

[17] R. Mohite and L. Ouarbya, "Interpretable anomaly detection: A hybrid approach using rule-based and machine learning techniques," Goldsmiths, Univ. London, London, UK, 2023.

[18] K. Adilakshmi, N. Sreelatha, D. Reshma, and J. Mounika, "A hybrid machine learning approach to real-time cyber threat detection and response for next-generation network security," *Power Syst. Technol.*, vol. 49, no. 4, 2025.

[19] R. Salles, B. Lange, R. Akbarinia, F. Masseglia, E. Ogasawara, and E. Pacitti, "Scalable and accurate online multivariate anomaly detection," *Inf. Syst.*, vol. 131, Jun. 2025, Art. no. 102524, doi: https://doi.org/10.1016/j.is.2025.102524.

[20] S. BitSight, "What is continuous cybersecurity monitoring?" Security Scorecard.com. https://securityscorecard.com/blog/what-is-continuous-cybersecurity-monitoring/ (updated June 23, 2025).

[21] M. U. Ashraf *et al.*, "A hybrid deep learning model for network intrusion detection from raw network bytes," *IEEE Access*, vol. 12, pp. 25544–25557, 2024.

[22] H. Alqahtani *et al.*, "A hybrid deep learning model for advanced persistent threat attack detection," in *Proc. 5th Int. Conf. Future Netw. Distrib. Syst. (FNDS)*, Dubai, UAE, 2021, pp. 92–

97, doi: https://doi.org/10.1145/3508072.3508085.

[23] Z. Chen, Z. Li, J. Huang, S. Liu, and H. Long, "An Effective Method for Anomaly Detection in Industrial Internet of Things Using XGBoost and LSTM," Sci. Rep., vol. 14, 2024, Art. no. 23969, doi: https://doi.org/10.1038/s41598-024-74822-6.

[24] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A Review of Novelty Detection," *Signal Proc*., vol. 99, pp. 215–249, Jun. 2014, doi: https://doi.org/10.1016/j.sigpro.2013.12.026.

[25] A. K. Singh and N. Kumar, "Comparative Analysis of LSTM, XGBoost and Hybrid Approaches in Credit Card Fraud Detection," 2024.

[26] E. Kahraman, B. O. Taiwo, S. Hosseini, Y. Fissha, V. A. Jebutu, and A. A. Akinlabi, "Blast toes volume estimation for post-blast efficiency: A comparative Analysis of hybrid ensemble learning, voting, and base AI-algorithms," *Res. Seq.*, Mar. 2024, doi: https://doi.org/10.21203/rs.3.rs-4014302/v1.

[27] E. Efatinasab, A. Brighente, D. Donadel, M. Conti, and M. Rampazzo, "Towards Robust Stability Prediction in Smart Grids: GAN-Based Approach Under Data Constraints and Adversarial Challenges," *Internet Things*, vol. 33, Sep. 2025, Art. no. 101662, doi: https://doi.org/10.1016/j.iot.2025.101662.