An Explainable Hybrid Machine Learning Framework for Network Threat Prediction

Niya Maryam[1], Shaik Asha[2]

[1]M.Tech, Department of Computer Science and Engineering, Muffakham Jah College of Engineering and Technology, Hyderabad, Telangana 500034, India
Email: 160424742014@mjcollege.ac.in

[2]Assistant Professor, Department of Computer Science and Engineering, Muffakham Jah College of Engineering and Technology, Hyderabad, Telangana 500034, India
Email: shaik.asha@mjcollege.ac.in

Corresponding Author: Niya Maryam

***Abstract*— With the growing sophistication and prevalence of cyber threats, an intelligent and proactive security approach is required to accurately detect and predict cyber attacks. This study introduces a complex AI-driven cyber attack prediction framework using the CICIDS2017 dataset, which integrates ML, DL, GAI, and XAI techniques. The preparation of the data is comprehensive, such as removing missing values, duplicate data, encoding labels, data normalisation and dimension reduction through PCA, to help improve the learning efficiency. We train and test multiple ML classifiers including DT, RF, ET classifier, logistic regression, gaussian naive bays, and a voting classifier of combination of RF, light gbm and xgboost as well as multiple DL models including DNN, CNN, LSTM, CNN–LSTM and CNN–LSTM–GRU. GAE, GAN and DistilGPT2 models are used to generate synthetic attack patterns and improve anomaly representation. The experimental results show that the Voting Classifier achieved the maximum accuracy of 99.6%, while LSTM achieved 99.3%, indicating that the VC and LSTM model had excellent detection accuracy for various attacks such as DoS, DDoS, PortScan, Bot and Infiltration attacks. For model interpretability use LIME and SHAP. The framework is developed in Flask which takes care of authentication, input processing, visualisation and categorisation of network traffic as benign or dangerous.**

"Keywords— Cyber attack prediction, Machine Learning, Deep Learning, Generative AI, Explainable AI, LSTM, Ensemble Learning, Intrusion Detection."

I. INTRODUCTION

The fast development of digital technologies and the broad interconnectedness of different industries have considerably increased the significance of strong cybersecurity procedures. The use of the internet by organisations, governments and individuals has grown rapidly, and so has the threat and sophistication of cyber attacks. Ransomware, phishing, DDoS attacks, data breaches, are a type of cyberattack that disrupts important services, results in major economic losses and sensitive information being leaked [1]. In an environment of evolving cyber threats, traditional defence strategies, which are mostly reactive, have been found to be ineffective, thus highlighting the need for intelligent, proactive and adaptive cyber security strategies [2].

AI and its allied fields such as ML, DL, NLP and GenAI are proving to be transformational tools in augmenting cybersecurity [3,4]. They offer predictive threat intelligence, real-time anomaly detection and automated mitigation tactics, moving cybersecurity from reactive defences to proactive, data-driven protection. ML models can be used to analyze traffic patterns on a network to identify very small deviations from normal, which could signify intrusion. It is possible to use NLP algorithms to identify and classify phishing attempts and malicious messages [3]. More advanced DL methods like CNN, LSTM networks and hybrid models of CNN-LSTM or CNN-Transformer have shown to be more accurate with the recognition of more complicated attack patterns in different cybersecurity applications [4]. Additionally, security systems can better withstand adaptive attacks by anticipating and preparing for attack routes that they have not yet seen thanks to the usage of generative AI models to simulate novel threat scenarios [5].

However, the issues of data quality, scalability, model interpretability, and integration with current infrastructure [6,7] still present. The rising adoption of XAI approaches, such SHAP and LIME, is intended to enhance the transparency and confidence in AI-based cybersecurity systems, allowing analysts to comprehend the reasoning underlying the automated predictions [8,9]. The incorporation of the above mentioned AI techniques into a practical and real-time application is essential for the development of an end-to-end cyber security solution to work on dynamic and high-volume networks [10].

The primary objective of this project is to design and develop an integrated AI-based cyber security system comprising of models like ML, DL, NLP, and GenAI for predictive threat detection and mitigation. Through preprocessing and dimensionality reduction, the system meets real-time inference to advance detection accuracy, optimize decision making and enhance digital resilience across various infrastructure, so as to provide proactive defense against new cyber threats.

II. RELATED WORK

As digital infrastructure becomes more and more vital to the operation of society, cybersecurity is increasingly being studied and AI is a powerful tool for predictive and adaptive cyber defence. Several researchers have tried to address the

issue of aligning cybersecurity frameworks with international standards like ISO/IEC 27000, 27001 and 27002 which provide a set of organised principles for information security management and risk mitigation [11]. These requirements highlight the necessity of methodical ways to safeguarding digital assets, as well as AI-based solutions for threat detection and prevention.

Several ML and DL algorithms have been applied to integrate AI with cybersecurity for threat detection and anomaly detection, and for predictive analytics [12]. The research results show that classic cybersecurity techniques, such as signature-based and heuristic techniques, cannot cope with the increased complexity of cyberattacks. The analysis of network traffic, malicious pattern identification and ZD detection have been demonstrated to be well-suited for supervised and unsupervised ML models [13]. The models are trained with large amounts of data, and can be more accurate and have fewer false alarms, allowing for more proactive defense.

AI may be a crucial part of cybersecurity in predictive analysis and anomaly identification. In classification of normal and abnormal network behaviour, the traditional ML methods include DT, RF and SVM, which can be used to detect threats early [14]. These are the key strategies of predictive cybersecurity, which can predict future attacks based on previous trends. In addition, hybrid techniques based on a combination of ML and DL architectures like as CNNs and LSTM networks might improve detection of complex threat signatures especially for large-scale IoT and enterprise networks [16,17].

In the realm of AI-powered cybersecurity, XAI is proving to be a key factor in tackling the problems of transparency and interpretability of black-box systems. The SHAP or LIME can be used to explain the model's decision and provide the most important features, and can help build confidence in threats' automated forecasts [15]. It also has been found that XAI models can be susceptible to adversarial black-box attacks that can distort explanations and compromise cybersecurity defenses [20]. These findings indicate that while XAI enhances ease of use and trust, further actions are needed to ensure its robustness against malicious interference.

Recently, there has been an increasing focus on anomaly detection in networks using DL, which is able to detect temporal correlations and complex patterns in the network traffic data [18]. Techniques like CNNs, LSTMs and hybrid CNN-LSTM models perform more accurately and are more scalable when it comes to detecting APT and botnet activities. When used with real-time data processing pipelines that enable detection and response to evolving cyber threats, these models are effective.

Cybersecurity threats in cloud systems have also been extensively researched since the use of cloud has provided many more attack fronts and data privacy concerns [19]. AI-powered monitoring and predictive analytics on cloud infrastructures have also boosted the resilience of cloud services by detecting anomalous events, policy violations, and unauthorized access. The combination of AI models with continuous learning mechanisms means that the defence systems can react to new and changing attack techniques while maintaining high efficiency and low latency.

III. Materials And Methods

The suggested system enhances cybersecurity by using predictive intelligence with a blend of ML (DT, RF, LR, NB, VC), DL (DNN, CNN, LSTM, CNN+LSTM, CNN+LSTM+GRU), and Generative AI models (VAE, GAN, DistilGPT2). The system is based on the data of the CICIDS2017 dataset and data preprocessing, normalisation and PCA dimensionality reduction are used to make the computational efficiency and feature relevance. LIME and SHAP are among the tools employed in Explainable AI to help users understand the importance of features and the reasoning behind AI decisions [22,25]. Subtle and changing attack patterns are identified by DL architectures like hybrid CNN and LSTM. An AI algorithm that generates unique threat scenarios, predicting intrusions proactively [21]. The system is deployed with a web interface using Flask, and the data can be entered in real time, and models can be inferred and visualised. The end-to-end approach offers an adaptable, interpretable and scalable cybersecurity solution suitable for addressing complex and evolving cyber threats.

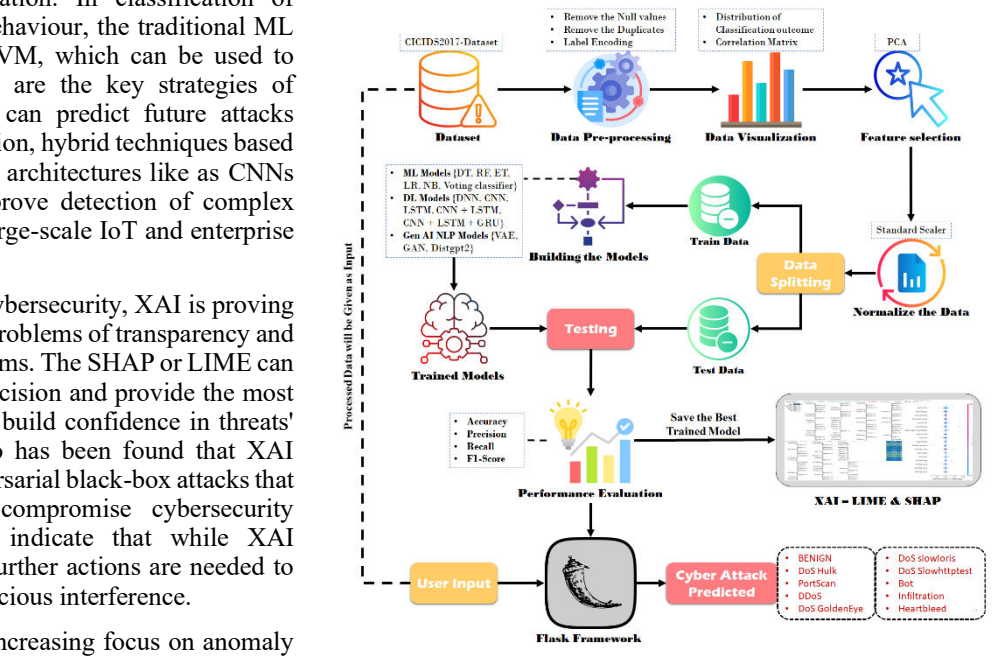

Fig. 1. System Architecture

In Figure 1, it shows that the methodology for the prediction of cyberattacks starts from the data CICIDS2017 pre-processing, visualisation and feature selection with PCA. Once the data is normalized and split, ML models and DL models are trained, evaluated and deployed with the help of the Flask framework for predicting the user's choices.

A) Dataset Collection:

This research uses CICIDS2017 dataset that provides a collection of complete network traffic captures for various cyber-attack scenarios including DoS, DDoS, PortScan, Botnet, Infiltration and Heartbleed and normal traffic. It contains 2,214,469 samples and 79 attributes like flow statistics, packet lengths, inter-arrival time, and flag counts, to study network behavior in detail. High quality and well labelled dataset with a diverse set of common and sophisticated attacks make it a good dataset to test predictive cybersecurity models. The fact that the number of benign

attacks and unusual attacks are significantly imbalanced makes it necessary to carefully preprocess and feature engineer the data for efficient model training. CICIDS2017 is an industry standard for research of the AI-based intrusion detection and anomaly detection field.

Table.1 Sample of the CICIDS2017 Dataset

S.No	Destination Port	Flow Duration	Total Fwd Packets	Total Bwd Packets	Total Length of Fwd Packets	Total Length of Bwd Packets	Fwd Packet Length Max	Fwd Packet Length Min	Fwd Packet Length Mean	Fwd Packet Length Std	…	Active Min	Idle Mean	Idle Std	Idle Max	Idle Min	Label
0	54865	3.0	2.0	0.0	12.0	0.0	6.0	6.0	6.0	0.0	…	0.0	0.0	0.0	0.0	0.0	BENIGN
1	55054	105.0	1.0	1.0	6.0	6.0	6.0	6.0	6.0	0.0	…	0.0	0.0	0.0	0.0	0.0	BENIGN
2	55055	52.0	1.0	1.0	6.0	6.0	6.0	6.0	6.0	0.0	…	0.0	0.0	0.0	0.0	0.0	BENIGN
3	46236	34.0	1.0	1.0	6.0	6.0	6.0	6.0	6.0	0.0	…	0.0	0.0	0.0	0.0	0.0	BENIGN
4	54863	3.0	2.0	0.0	12.0	0.0	6.0	6.0	6.0	0.0	…	0.0	0.0	0.0	0.0	0.0	BENIGN

B) Pre-processing:

Preprocessing of data is a key step in preparing raw network traffic data for AI-based cybersecurity research. It cleans data for quality, consistency and dependability by filling in missing values, removing duplicate values and transforming categorical labels into a numerical format. Well done, if the model is well preprocessed, it will be more accurate and less computing load. Also, it enables the robust identification of subtle and developing patterns of cyber-attacks.

i) Data Preprocessing: The data processing pipeline begins with removing missing data since it can negatively impact the model's accuracy. Any data that are redundant or biased are eliminated. Label encoding technique is employed to convert categorical attack categories to numeric values, which can be used to train ML and DL algorithms. Features further are normalised and scaled to make their amplitude closer to the normalised amplitude of all features to ensure convergence and stability during training. The preprocessing pipeline can effectively cleanse and structure the data, providing high-quality data inputs for AI models to facilitate accurate and timely detection and classification of cyber threats.

ii) Data Visualization: This shows data trends and the type of cyber-attack, that is common on the network. The categorisation results can be visualised and provide insight into the class imbalance and the abundance of benign traffic in comparison to rare attack types, such as Infiltration or Heartbleed. Correlation matrices are created to determine correlations and dependencies between characteristics, which can help in detecting multicollinearity and duplicated qualities. Graphical information such as histograms, box plots, and heat maps. These visual insights help with feature engineering, selection, and model building, making sure predictive algorithms focus on the most relevant features for accurate and efficient threat identification.

iii) Feature Selection: Feature selection is used to minimise dimensionality, improve computing efficiency and interpretability of the model. We use PCA to convert the original high-dimensional feature space to a lower-dimensional space of uncorrelated principal components, which include much of the information of the original space. PCA is used to remove redundant and noisy data by retaining those variables with the most variation in the data. This dimensionality reduces training time, reduces overfitting and improves the model's generalisability to new data. Feature selection can be applied to AI models to identify the most relevant features, leading to more powerful, scalable, and accurate predictions in cybersecurity.

iv) Normalization: Normalisation is an important pre-processing step to standardise the size of numerical features, such that all input variables contribute equally to the model training. In this study, we use Standard Scaler normalisation to make all features have mean = 0 and stddev = 1. Especially for algorithms which use the magnitude of features like neural networks or distance-based classifiers, this is important because it avoids the larger the feature the more influence the learning process would have. Normalisation helps in achieving the better convergence during training as it stabilises the gradients, which reduces the chances of the gradients bursting or vanishing. Normalisation is a technique that scales the dataset evenly to boost the performance of the model, provide strong learning and increase the accuracy and stability of the AI-based cyber-attack detection systems.

C) Train & Test:

Data splitting lies at the heart of building and evaluating cybersecurity prediction models. CICIDS2017 consists of 80% training samples and 20% testing samples, typically. With this, algorithms learn patterns and also remember the information they haven't seen yet for objective review. Good separation ensures that the learnt behavior is good with the learned inputs but would be bad with new ones (which is called overfitting). Stratified splitting can be used to keep the class proportions in splits the same if the uncommon attack categories are rarely used in practice and benign traffic is the vast majority of the time.

D) Algorithms:

DT it is a decision-representation in the form of a tree. It divides the data based on feature values and makes

predictions based on the divisions. It categorizes the network traffic as normal or malicious, and offers interpretability and insights into some of the most important factors which influence the attack, in the context of cyber attack prediction.

$$I\,(i) = 1 - \sum_{i=1}^{k} p_i^2 \qquad (1)$$

RF builds numerous decision trees from random subsets of features, and fuses their predictions into a firm prediction. In the field of cybersecurity, it is employed for handling high-dimensional data, boosting precision, reducing overfitting, and uncovering meaningful features to ensure accurate intrusion detection.

The Gini Equation given below:

$$Gini = 1 - \sum_{i=1}^{C} (Pi)^2 \qquad (2)$$

In ensemble learning, ET randomises splits in trees. They are excellent at pattern recognition, dealing with unbalanced or noisy data, and predicting more quickly and accurately in modelling complex relations between components in cyber threat detection.

LR is applied in detecting network anomalies in a cyberattack to predict the probability of belonging to a class in which it has detected malicious network activity. It provides probabilities, interpretability, simplicity and a foundation for cybersecurity analytics performance benchmarking.

The equation shows the logistic sigmoid function prediction of class probabilities.

$$\hat{y}_i = \sigma(w^T x + b) = \frac{1}{1 + e^{-(w^T x + b)}} \qquad (3)$$

Naïve Bayes a classification algorithm based on Bayes' Theorem with assumption of independence of features. It provides fast, low cost and effective traffic classification in cyberattack prediction. It can be used for large data set and host real-time intrusion detection.

Bayes Theorem formula is given as:

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \qquad (4)$$

Voting Classifier is an ensemble of classifiers technique that is designed to achieve the highest accuracy and minimum bias by fusing the results of multiple classifiers. It complements the precision and recall in the cybersecurity domain by capitalizing on the strengths of the tree-based learners in stable and reliable detection of normal and malicious network traffic.

In learning, DNNs have several layers that learn complex and non-linear patterns in the data. They are able to gather high level information from network data to predict cyber attacks, making them more accurate and able to detect more sophisticated attacks than traditional algorithms.

CNNs extract the spatial features through convolutional and pooling layers. In cyber attack detection they detect local correlations in the network data, avoid manual feature engineering, and effectively identify hidden patterns and anomalies in high-dimensional cybersecurity data.

The fundamental functioning of CNNs.

$$S(i,j) = \sum_{m} \sum_{n} I(i+m, j+n) \cdot K(m,n) \qquad (5)$$

LSTMs are used to model long-term dependencies in sequential data, using memory cells and gates. They gather temporal patterns from a network traffic to predict cyber attacks to improve accuracy and to identify emerging threats on time-series behaviors.

The hidden state update of LSTM is hidden in the following equation.

$$h_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \cdot \tanh(C_t) \qquad (6)$$

This is a CNN-LSTM hybrid model that extracts spatial features from CNN and models temporal features from LSTM. It enhances the ability to recognize patterns and subsequent connections at local levels, resulting in better accuracy and understanding of complex cyber risks.

$$y = Softmax\left(W_o.LSTM\left(MaxPool\left(f(W_c.X_{i:i+k-1} + b_c)\right)\right) + b_o\right) \qquad (7)$$

The hybrid CNN, LSTM, and GRU model has been shown to be an effective way of capturing spatial and temporal patterns. GRU delivers accurate, flexible identification of complex, evolving cyber threats in network data, and simplifies the process without compromising on speed.

VAE reconstructs the learning distributions after learning the widely distributed data representations. It increases resilience and adjusts models to unexpected threats by aiding in anomaly detection and modelling new cybersecurity attack patterns.

GANs consist of a generator and a discriminator, and are trained adversarially. They generate realistic attack scenarios for training to detect attacks that are not expected or are zero-day attacks, and to strengthen adaptive threat intelligence.

DistilGpt2 is a small model for text and sequence data based on the transformer architecture. It is used in cybersecurity to analyse logs, alerts and reports in order to detect intrusion patterns, for rapid, scalable and context-aware threat intelligence for real-time anomaly detection.

E) Integration of XAI & Flask:

Suggested Cybersecurity System: For enhancing the transparency and interpretability of the model predictions, we introduce XAI using LIME and SHAP in the suggested cybersecurity system. LIME provides local approximation of complex model behaviour by showing the contribution of features for specific predictions and SHAP measures the influence of each feature over the entire dataset. This pairing enables security analysts to gain insight into the reasons behind a particular network traffic being classified as malicious or benign, and it ensures that AI-powered intrusion detection systems have confidence and accountability.

We use the Flask framework to provide an interface on the web for interacting in real-time with the AI models. It enables

user to input real-time network data, activate model inference and visualize the results using XAI-generated explanations. The lightweight and modular architecture of Flask facilitates seamless deployment, quick prototyping and seamless integration of pre-processing, predictive analytics, and interpretability features into an end-to-end cybersecurity platform.

IV. Experimental Results

Accuracy: The accuracy of a test is its capacity to appropriately separate the sick and healthy cases. The proportion of true positive and true negative cases divided by the number of cases assessed is an estimate of the accuracy of the test. This can be written as a mathematical equation:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (8)$$

Precision: is the fraction of accurately categorised positive instances or samples to the total instances or samples classified as positives. We can then calculate the precision as follows:

$$Precision = \frac{True\ Positive}{True\ Positive\ + False\ Positive} \quad (9)$$

Recall: Recall is a measure of ML that measures how well a model is able to identify all of the examples within a particular class. It is the proportion of true predictions of positive observations to overall number of observations in real class. Provides a sense of the proportion of cases being included in the model's estimates.

$$Recall = \frac{TP}{TP\ +\ FN} \quad (10)$$

F1-Score: In ML, F1 score is one of the metrics used for the evaluation of the accuracy of a model. It is a combination of the precision and recall ratings of a model. Accuracy statistic is the number of accurate predictions made by the model on the entire set of data.

$$F1\ Score = 2 * \frac{Recall\ X\ Precision}{Recall + Precision} * 100 \quad (11)$$

Table.2 Performance Evaluation

ML Model	Accuracy	Precision	Recall	F1-Score
DT	0.996	0.996	0.996	0.996
RF	0.996	0.996	0.996	0.996
ETC	0.996	0.996	0.996	0.996
LR	0.907	0.908	0.907	0.904
GNB	0.304	0.900	0.304	0.382
Voting (RF + LGBM + XGB)	0.996	0.996	0.996	0.996
DNN	0.980	0.980	0.980	0.980
CNN	0.980	0.981	0.980	0.980
LSTM	0.993	0.993	0.993	0.993
CNN + LSTM	0.990	0.990	0.990	0.990
CNN + LSTM + GRU	0.981	0.982	0.981	0.981

Table 1 shows the performance of all the models (ML and DL) with accuracy, precision, recall and F1 score.

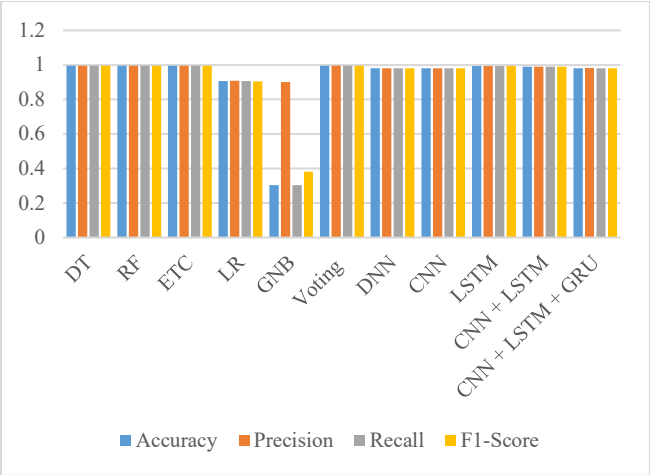

Fig.3 Comparison Graph

Figure 3 is a bar chart which compares Accuracy, Precision, Recall and F1-Score of several categorisation models.

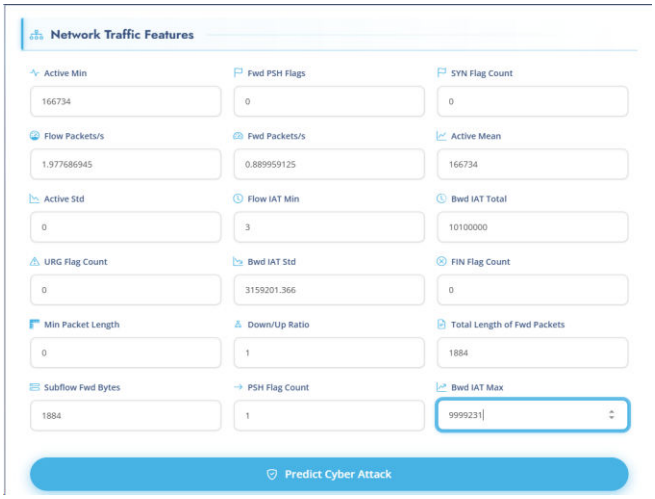

Fig.4 Upload Input Data

The input screen of this Fig. 4 is for submission of values of features of network traffic for final cyber-attack prediction analysis.

Fig.5 Predict Result

The final output result is “BENIGN” (SAFE), with a confidence score of “100.0%”, and regular network traffic, as shown in Figure 5.

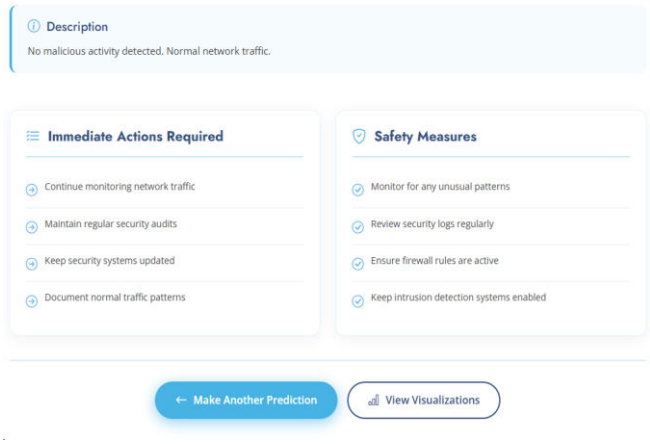

Fig.6 Immediate Actions & Safety Measures

This Figure 6 depicts the events that are Immediate Actions Required (e.g., monitoring, audits) and Safety Measures (e.g., firewall rules, intrusion detection) that will be required for the anticipated network status.

V. Conclusion

The proposed AI-based cybersecurity system efficiently combines high-level algorithmic intelligence with realistic cyber threat detection and prediction. Excellent detection accuracy and interpretability were accomplished by the system by systematic pre-processing, feature optimisation using PCA, and model evaluation in the context of ML, DL, and Generative AI paradigms. The ensemble voting classifier of the ML models (RF, LightGBM and XGBoost) performed best, achieving an accuracy of 99.6%, a precision of 99.6% and a recall of 99.6%, which means that the model has a good generalisation ability and is robust in distinguishing the different attack categories. The accuracy of 99.3% obtained by the LSTM model was the highest, which was able to model the temporal dependencies and complex patterns of incursion. Generative AI models such as VAE, GAN and DistilGPT2 have been added to enhance the ability to simulate a synthetic attack scenario to feed adaptive cybersecurity intelligence. LIME and SHAP were used in combination to achieve transparency using explainable AI approaches, which visualised feature contributions and model reasoning, making predictions interpretable and trustworthy. This overall result indicates that ensemble-based ML models when integrated with sequential DL architectures and interpretability frameworks is a very effective, explainable and scalable approach for automated current cyber attacks prediction and defense.

Future directions involve development of the system into autonomous, adaptive cybersecurity systems that are able to adapt to dynamic threats in real time. Ongoing self-updating and resilience is made possible by using online learning alongside reinforcement learning. Threats can be better understood in context using transformer designs such as BERT and GPT. The more the amount of data sources there are are in different fields, the better the model will be able to generalize. Cloud and Edge deployment of Scalability and Latency Blockchain for Log Management and Federated Learning enhances privacy, integrity and collaborative defence in global dispersed cyber infrastructure networks and ecosystems.

References

[1] Ajala, O. A., Okoye, C. C., Ofodile, O. C., et al. (2024). Review of AI and machine learning applications to predict and Thwart cyber-attacks in real-time. Magna Scientia Advanced Research and Reviews, 10(1), 312-320.

[2] Nadella, G. S., & Gonaygunta, H. (2024). Enhancing cybersecurity with artificial intelligence: Predictive techniques and challenges in the age of IoT. International journal of science and engineering applications, 13(04), 30-33.

[3] Arif, A., Khan, M. I., Khan, A. R. A., et al. (2025). AI-Driven Cybersecurity Predictions: Safeguarding California's Digital Landscape. International Journal of Innovative Research in Computer Science and Technology, 13, 74-78.

[4] Raza, A., Ali, A. K. S., & Hussain, A. A. (2024). AI-DRIVEN APPROACHES TO CYBER AND INFORMATION SECURITY: MACHINE LEARNING ALGORITHMS FOR THREAT PREDICTION AND ANOMALY DETECTION. Spectrum of Engineering Sciences, 2(4), 565-573.

[5] Rahman, M. K., Dalim, H. M., & Hossain, M. S. (2023). AI-Powered solutions for enhancing national cybersecurity: predictive analytics and threat mitigation. International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence, 14(1), 1036-1069.

[6] Kao, D.-Y., Hsiao, S.-C., & Tso, R. (2019). Analyzing WannaCry ransomware considering the weapons and exploits. In Proc. 21st Int. Conf. Adv. Commun. Technol. (ICACT), pp. 1098–1107.

[7] Bresniker, K., Gavrilovska, A., Holt, J., et al. (2019). Grand challenge: Applying artificial intelligence and machine learning to cybersecurity. Computer, 52(12), 45–52.

[8] Husák, M., Komárková, J., Bou-Harb, E., & Celeda, P. (2019). Survey of attack projection, prediction, and forecasting in cyber security. IEEE Commun. Surveys Tuts., 21(1), 640–660.

[9] Mohamed, N. (2023). Current trends in AI and ML for cybersecurity: A state of-the-art survey. Cogent Eng., 10(2), 1–11.

[10] Chan, L., Morgan, I., Simon, H., et al. (2019). Survey of AI in cybersecurity for information technology management. In Proc. IEEE Technol. Eng. Manage. Conf. (TEMSCON), pp. 1–8.

[11] Disterer, G. (2013). ISO/IEC 27000, 27001 and 27002 for information security management. J. Inf. Secur., 4(2), 92–100.

[12] Li, J. H. (2018). Cyber security meets artificial intelligence: A survey. Frontiers Inf. Technol. Electron. Eng., 19(12), 1462–1474.

[13] Sarker, I. H., Kayes, A. S. M., Badsha, S., et al. (2020). Cybersecurity data science: An overview from machine learning perspective. J. Big Data., 7, 1–29.

[14] Aggarwal, D., Sharma, D., & Saxena, A. B. (2023). Role of AI in cybersecurity through anomaly detection and predictive analysis. J. Informat. Educ. Res., 3(2), 1–12.

[15] Srivastava, G., Jhaveri, R. H., Bhattacharya, S., et al. (2022). XAI for cybersecurity: State of the art, challenges, open issues and future directions. arXiv:2206.03585.

[16] Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). AI-driven cybersecurity: An overview, security intelligence modeling and research directions. Social Netw. Comput. Sci., 2(3), 1–18.

[17] Ucci, D., Aniello, L., & Baldoni, R. (2019). Survey of machine learning techniques for malware analysis. Comput. Secur., 81, 123–147.

[18] Kwon, D., Kim, H., Kim, J., et al. (2019). A survey of deep learning-based network anomaly detection. Cluster Comput., 22(S1), 949–961.

[19] Nafea, R. A., & Almaiah, M. A. (2021). Cyber security threats in cloud: Literature review. In Proc. Int. Conf. Inf. Technol. (ICIT), pp. 779–786.

[20] Kuppa, A., & Le-Khac, N.-A. (2020). Blackbox attacks on explainable artificial intelligence(XAI) methods in cyber security. In Proc. Int. Joint Conf. Neural Netw. (IJCNN), pp. 1–8.

[21] Ahmed, K. D., & Askar, S. (2021). Deep learning models for cyber security in IoT networks: A review. Int. J. Sci. Bus., 5(3), 61–70.

[22] Gerlings, J., Shollo, A., & Constantiou, I. (2020). Reviewing the need for explainable artificial intelligence (xAI). arXiv:2012.01007.

[23] Jaswal, G., Kanhangad, V., & Ramachandra, R. (2021). AI and Deep Learning in Biometric Security: Trends Potential and Challenges. Boca Raton, FL, USA: CRCPress.

[24] Rudin, C. (2018). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. arXiv:1811.10154.

[25] Zhang, Z., Hamadi, H. A., Damiani, E., et al. (2022). Explainable artificial intelligence applications in cyber security: State of-the-Art in research. IEEE Access, 10, 93104–93139.

[26] Bandi, A., Adapa, P. V. S. R., & Kuchi, Y. E. V. P. K. (2023). The power of generative AI: A review of requirements, models, input–output formats, evaluation metrics, and challenges. Future Internet, 15(8).

[27] Babcock, J., & Bali, R. (2021). Generative AI With Python and Tensorflow 2: Create Images, Text, and Music With VAEs, GANs, LSTMs, Transformer Models. Birmingham, U.K.: Packt.

[28] Hitaj, B., Gasti, P., Ateniese, G., & Perez-Cruz, F. (2017). PassGAN: A deep learning approach for password guessing. arXiv:1709.00440.