International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal

Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Multi-Level Ensemble and Transfer Learning Framework for Cybersecurity Anomaly Detection

[1]Pratibha Pandey

Research Scholar, Department of Computer Science & Engineering, Maharana Pratap College of Technology, Gwalior

Pratibhapandey772@gmail.com

[2]Prof. Anjali Saxena

Professor, Department of Computer Science & Engineering, Maharana Pratap College of Technology, Gwalior

anjalisaxena11@gmail.com

Abstract- The swift increase in cyber threats in contemporary networked environments has led to the need of developing powerful and smart anomaly detectors. The present paper focuses on proposing a multi-level ensemble and transfer learning-based architecture to effectively detect anomalies in cybersecurity systems. Its main goal is to enhance the accuracy of detection and generalization to a wide range of data sets using hybrid methods of deep learning. This approach combines the results of two benchmark datasets, UNSW-NB15 and CICIDS-2017, and then preprocesses the data such as data cleaning and the treatment of missing values and feature selection using Select KBest. Exploratory Data Analysis would be conducted to learn the characteristics of data and the division of data and normalization would be done to make sure that the model is trained effectively. The suggested framework integrates three models: a dense neural network with transfer learning to learn more features, an LSTM model to learn temporal features, and an auto encoder to detect anomalies unsupervised. A combination of these models is done through a weighted ensemble method to make final predictions. Experimental analysis shows that the proposed model obtains a precision of 82.68 on the UNSW dataset and 96.29 on the CIC dataset. The findings indicate the efficiency of the ensemble method to enhance the performance of the anomaly detection and flexibility to various cybersecurity conditions.

Keywords- Anomaly Detection, Cybersecurity, Ensemble Learning, Transfer Learning and Deep Learning

I. Introduction

Over the past few years, the cybersecurity landscapes have become highly complicated and vulnerable due to the tremendous growth of digital infrastructures, cloud computing, Internet of Things (IoT), and Cyber-Physical Systems (CPS). Due to the growing integration of systems in organizations, the occurrence and complexity of cyber threats like malware, intrusion attempts,

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

distributed denial-of-service (DDoS) attacks, and insider threats have multiplied manifold [1], [2]. The conventional security solutions such as rule-based security and signature based intrusion detection systems (IDS) are no longer adequate to deal with the dynamic and changes nature of the contemporary cyberattacks. Such traditional methods are usually ineffective in identifying zero-day attacks and never-before-seen anomalies since they compare them with predefined patterns. In order to address these shortcomings, machine learning (ML) and deep learning (DL) have become effective methods of cybersecurity system anomaly detection. The techniques allow systems to acquire patterns based on large-scale data and detect deviations that can be a signal of malicious behavior. Nevertheless, single-model methods are usually limited by overfitting and lack of generalization and sensitivity to noise in high-dimensional data. Also, labeled cybersecurity datasets are usually scarce, lopsided, or domain-biased, further constraining the capabilities of standalone models [3], [4].

Ensemble learning in this regard has been of interest as a promising method to enhance the detection accuracy and robustness through the combination of many models [5], [6]. Ensemble techniques use the strengths of the various algorithms to achieve more reliable predictions in comparison to single models. Simultaneously, transfer learning has become an attractive method to overcome the problem of data scarcity and domain adaptation through the utilization of pretrained models on large datasets and their application to related tasks [7]. A combination of ensemble learning and transfer learning offers a potent scheme in constructing scalable and adaptive anomaly detectors [8].

This research paper introduces a multi-level ensemble and transfer learning system that can be used to improve the performance of anomaly detectors in cyber security systems. The suggested framework takes advantage of various base learners at various stages and synthesizes transfer learning processes to enhance feature extraction and model generalization. The system seeks to identify both known and unknown patterns of attack and minimize the false positives and computational cost by integrating heterogeneous models and embedding pre-trained deep learning structures.

The multi-level architecture enables the architecture to work in a hierarchical step, with the first level models extracting features and making coarse predictions, and then higher-level ensemble models that narrow down the predictions. Such a stratified method enhances the process of decision-making by synthesizing the results/output of more than one perspective, which enhances the reliability and applicability of the system [9], [10], [11], [12], [13]. Moreover, transfer learning allows the model to take advantage of already learned representations, thus it does not require large labeled datasets and shortens the training time [14], [15].

The other important feature of the proposed approach is that it can be adjusted to the real-time cybersecurity settings. Contemporary networks are capable of producing vast amounts of

information at any given moment, making anomaly detection systems to be efficient and scalable. Combination of optimized ensemble strategies with transfer learning is such that the model is able to process high data stream with high detection performance. Also, the architecture can support various kinds of data, such as network traffic, system logs, and user behavior trends, which is why it can be used in various cybersecurity applications, and in this research, the gap between the old approaches to anomaly detection and the new intelligent systems is going to be bridged by creating a hybrid framework that will combine the advantages of ensemble learning and transfer learning. This proposal will solve major problems in the data imbalance, complexity of features, and generalization of models, which will help create more resilient and secure cybersecurity mechanisms.

Contributions of the Study

The key contributions of this research are as follows:

1. **Development of a Multi-Level Ensemble Architecture:** A hierarchical ensemble model aims to integrate various machine learning and deep learning models to enhance anomaly detection.
2. **Integration of Transfer Learning:** The paper uses the pre-trained models to improve feature extraction and lessen reliance on high-labeled datasets.
3. **Hybrid Detection Mechanism:** The proposed system is a combination of the supervised and deep learning-based method to identify the complex and unknown pattern of attack.
4. **Improved Generalization and Robustness:** The model reduces overfitting and enhances performance on a variety of cybersecurity data by utilizing ensemble strategies.
5. **Scalability for Real-Time Applications:** The architecture is made in such a manner that it is capable of dealing with real-time and large scale data streams in contemporary settings.

This is novel research since it combines multi-level ensemble learning with transfer learning and puts them together in a single architecture to detect cybersecurity anomalies. In contrast to the traditional single-layer or standalone models, the proposed framework includes a hierarchical decision-making process, where several base learners and meta-learners are involved to refine the predictions, the transfer learning is utilized in the layers of the ensemble, which allows the system to use the prior knowledge to more effectively represent the features, especially in cases of limited or imbalanced data. The architecture also focuses on adaptive learning and modular design enabling easy integration of various models and datasets.

This set of hierarchical ensemble modeling, knowledge transfer across domains, and scalability of the design make the proposed approach unique compared to the current ones and offers a powerful solution to identifying advanced cyber threats in dynamic settings.

II. Literature Review

Aziz et al., 2025[16] solves the issue of more advanced breach approaches that remain ahead of conventional safety approaches. The importance of data security and privacy as an area of research

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

was significantly high in recent times where privacy protection emerged as an important aspect in the creation of artificial intelligence-oriented IOTs. The End-to-End VPN security has an essential role especially in connecting smart objects in the Internet of Things (IoT) environments. It noted that security is a crucial issue in the End-to-End data security approach. The solution to this is the use of the Machine Learning (ML) algorithms on a popular intrusion detection systems (IDS) to ensure that a more advanced cybersecurity assault can be addressed with an efficient and effective detection system. The present paper gives a detailed discussion of the Virtual Private Network (VPN) technologies with particular focus on their significance in the current networking environment in terms of facilitating secure communications across untrusted networks such as the internet. VPNs have also developed greatly to meet the increased requirements of data security both in individual and business settings. This paper explores the different data security protocols like PPTP, L2TP/IPsec, OpenVPN, IKEv2/IPsec, and WireGuard, analyzing their security features, advantages, and disadvantages. Additional challenges surrounding VPNs emerging in the paper are the sophisticated cyber threats, and the effects of the emerging technologies, such as quantum computing. Moreover, the work provides future directions, including implementing AI to detect threats dynamically and creating quantum-resistant VPN protocols. Through this analysis, the aim is to provide actionable insights into optimizing VPN usage for enhanced network security in an increasingly complex digital landscape.

Ahmadi-Assalemi et al., 2025[17] present a new anomalous behaviour detection model of ICS data streams of physical plant sensors. One-class classification model is created, with stream rebalancing, and then adaptive machine learning algorithms that may include drift detection techniques to find anomalies in physical plant sensor data. On ICS datasets, the approach of the authors is demonstrated. Further, a use case is provided to demonstrate the applicability of the model to post-incident investigations as a defence-in-depth capability in ICS. The results of the experiment indicate that the proposed model has an overall Matthews Correlation Coefficient score of 0.999 and a Cohen Kappa score of 0.9986 on single type anomalous behaviour limited to a single data stream. The performance on wide data streams is 0.981 MCC score and 0.9808 K score in the prevalence of various forms of anomalous instances.

Dayarathne et al., 2025[18] emphasizes the application of deep learning models, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, to real-time grid data anomaly detection. These models that are implemented based on a PSCADsimulated dataset that is supplemented with artificial cyber-attacks display significant improvements in detection and prevention of threats. The paper highlights the challenges in defining cyber risks in grids with high renewable penetration, including frequency instability and lack of system inertia, and proposes energy storage and advanced forecasting options to address them. The proposed models attain more than 98% accuracy in the detection of cyber threats by

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

integrating a new pre-processing model based on feature derivatives, which offer a solid framework to safeguard smart power grids against the emerging cyber threats.

Wong et al., 2025[19] introduced an Anomaly Detection using Machine Learning Algorithms (AD-MLA) model according to which a Random Forest model is used to precisely identify abnormal activities. The framework contains feature selection, data clustering, and an energyefficient routing method that takes into account the node energy, signal strength, hop count, and link stability. Assessment of AD-MLA indicates that it minimizes the false alarms, improves the detection rates, and requires less power and less computing energy. It provides an intelligent, fast and effective security system in real time VANET environment and this makes it suitable to transportation systems where high reliability and safety are the order of the day. The proposed framework combines a Random-Forest-based anomaly detector with smart feature selection and a routing approach that is energy-efficient and takes into consideration the residual energy, signal strength and stability of the links to systematically tackle these problems. This method has 95.33% accuracy, 96.09% recall, 94.25% computational efficiency, and 91.45% resource-use efficiency. This is an effective solution to the scalability, latency and energy challenges that the past systems have experienced when integrating the blockchain technology and deep learning architectures. Sikder, 2025[20] present new architecture that combines OD methods with AI confidence and situational based approaches in CPSs; create- ing various pipelines and DL models that improve anomaly detections, forecasting precision, and proactive decision support in WDSs and APSs; and show- casing quantifiable cybersecurity, operational performance, and predictive performance improvements with real-world and synthetic data. All these attempts bring about more trustful and sustainable CPSs. The results of ex-periments are documented, analyzed and presented and it is seen that these contributions fill the gap between complicated theoretical models and practical applications in the real world. **TABLE I**. LITERATURE SUMMARY

Author & Year	**Method / Approach**	**Key Results**	**Research Gap**
Mavikumbure et al. (2024) [21]	Proposed a ResNet Autoencoder-based Cyber-Physical Anomaly Detection System using CAN protocol data	Achieved highest accuracy compared to OCSVM, VAE, LSTM-AE, IForest, PCA, and LOF; low training and inference time with effective anomaly detection	Lacks integration with hybrid or ensemble models; limited exploration of combining supervised and unsupervised techniques for enhanced detection

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Gaba et al. (2024) [22]	Review of deep learning-based attack detection methods for Cyber-Physical Systems	Identified strong performance of deep learning models due to layered architecture and feature extraction capabilities	Does not propose a concrete model; lacks implementation of hybrid or ensemble frameworks and real-time validation
Nabi et al. (2024) [23]	Quantitative study using ML models (Decision Tree, Random Forest, SVM) with survey data analysis	68% adoption of ML; identified issues such as high false positives, data imbalance, and difficulty in detecting zero-day attacks	Focuses on perceptionbased analysis rather than model development; lacks advanced DL or ensemble-based anomaly detection approaches
Bajic et al. (2024) [24]	Proposed Human-Cyber-Physical System (HCPS) framework using big data analytics and edge computing	Demonstrated improved real-time anomaly detection and integration of human expertise in decisionmaking	Limited focus on automated intelligent anomaly detection models; lacks deep learning or ensemblebased optimization
Manias et al. (2024) [25]	Comprehensive review of Smart Grid cyber-physical security using ML, data-driven, and model-based approaches	Identified current trends, challenges, and emerging cyber threats in smart grids	Does not provide a unified detection framework; lacks implementation of advanced hybrid architectures like transfer learning and ensemble models

III. Methodology

The data-driven and deep learning approaches are incorporated in the proposed methodology to detect anomalies in cybersecurity systems. It starts by collecting the data using UNSW-NB15 and CICIDS-2017 data, then proceeds to preprocess data, which includes cleaning, missing values, and changing labels to binary ones. EDA is performed to comprehend the data distribution, correlations and outliers. Select K Best feature selection method also minimizes the dimensionality by selecting the most important attributes. The data is further divided into training and testing sets and

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

normalized using Standard Scaler. Lastly, a hybrid framework that integrates Dense (transfer learning), LSTM, and Auto encoder is applied, and they are combined with the use of a weighted ensemble method to make their final decision.

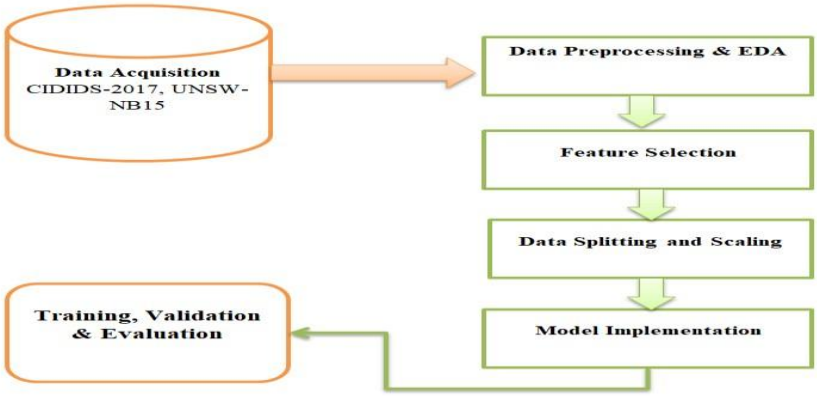

Fig. 1 Proposed Flowchart

A. **Data Collection**

The proposed approach to data collection of this research will allow the study to be robust and generalize the proposed anomaly detection model using various benchmark cybersecurity data sets. Two widely recognized datasets, namely UNSW-NB15 https://research.unsw.edu.au/projects/unsw-nb15-dataset and CICIDS-2017 https://www.unb.ca/cic/datasets/ids-2017.html , are employed in the process of capturing different normal and malicious network behaviors. These datasets are read in structured directories with several CSV files. It has a custom data loading functionality that is used to load all the available files in a systematic way, manage any encoding problem, and to fuse all the files into one database. When necessary to make them consistent, the label column is re-named to a common identifier (label) and categorical labels (like normal, benign) are converted into binary data (0 for normal traffic and 1 for anomalous activity) to allow a common binary classification scheme. The datasets usually have either missing/infinite values because of the real-world network situation; therefore, they are substituted and addressed accordingly.

One important process in the data collection stage is balancing of datasets. Cybersecurity datasets are usually not balanced, therefore, random sampling is employed to make normal and attack samples equivalent. This balanced dataset enhances the ability of models to learn and avoid bias with majority classes.

B. **Data Preprocessing**

Preprocessing of data is an important step that helps to enhance the quality of the data and train the model. First, duplicate records are eliminated in order to eliminate redundancy and make the model not repeat the same patterns. The dataset is also refined by substituting the infinite values with the

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

missing values and addressing them with the methods of statistical imputation. In particular, the gaps in the numerical features are occupied with the mean of respective columns and the rest of the gaps are occupied with zero to assure numerical stability.

Preprocessing stage is one that divides the data into the feature variables (X) and the target labels (y), with the label column corresponding to normal and abnormal classes. As machine learning models need numerical input, only numeric features (int and float types) are chosen, which means that it is compatible with the next steps of processing, and feature consistency is achieved by transforming column names into string form. This measure prevents possible errors in the training and selection of features, during the training of the model. The preprocessing pipeline guarantees that the dataset is clean, consistent and has no irregularities like missing or non-numeric values, this step improves the reliability of the data, mitigates noise, and makes the dataset in the most optimal form to be used in feature selection and model implementation.

C. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is carried out to learn more about the structure, distribution and quality of the dataset prior to model development. It involves examining balance of classes, relationships among features and existence of outliers. The EDA assists in the determination of patterns, anomalies, and pre-processing directions to make sure that the data is appropriate in constructing correct and dependable models.

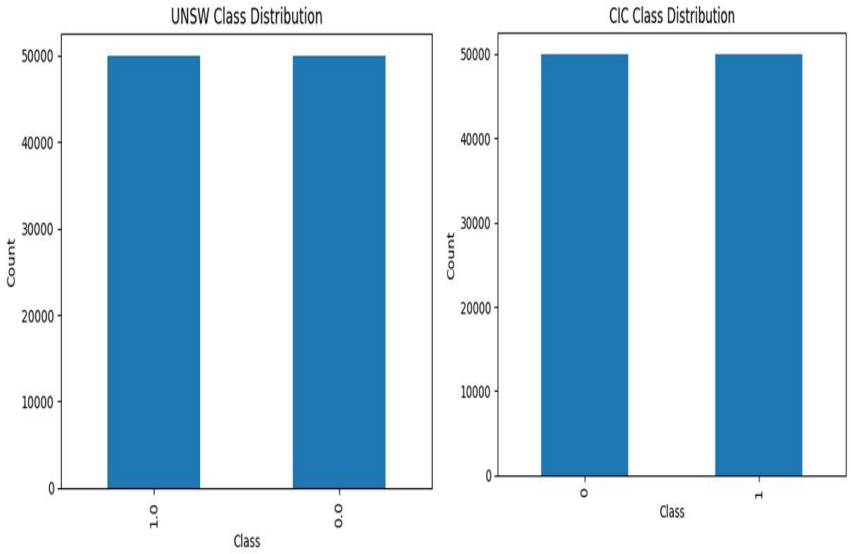

Fig. 2 Class Distribution of Both Datasets UNSW and CIC

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

The graphs of the distributions of classes indicate that there are equal numbers of normal and attack classes, which provide an unbiased model training and evaluation.

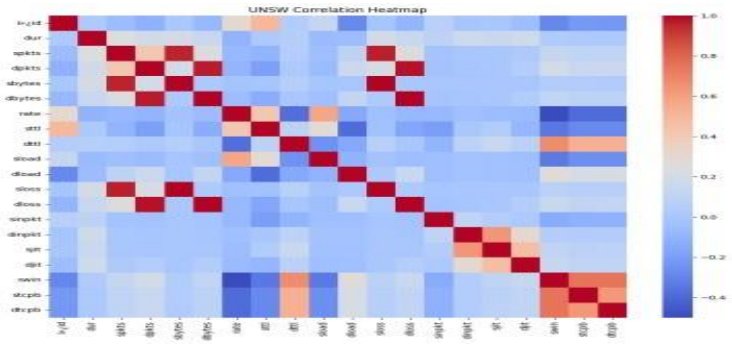

Fig. 3 Correlation Heatmap

The heatmap shows the correlations between features, showing strong positive and negative, which can be used to determine features that are important and eliminate redundancy.

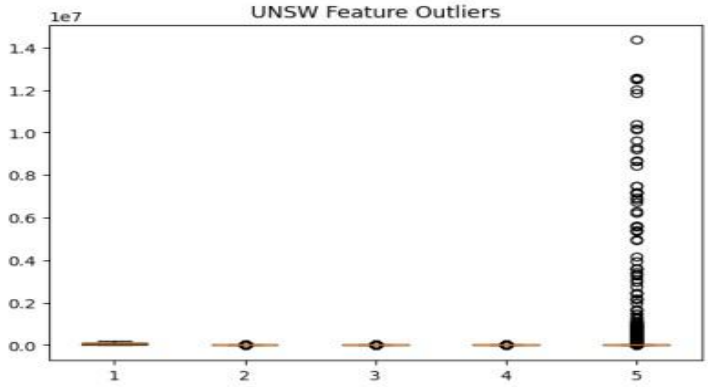

Fig. 4 Boxplot

Outlier plot shows that some of the features have extreme values and this reflects variability in data and preprocessing is required to enhance robustness and stability of the model.

D. **Feature Selection**

The feature selection is conducted to minimize the dimensionality and only the most relevant attributes that can be useful in detecting anomalies are kept. In our research, we used a statistical method of feature selection, Select K Best with ANOVA F-test (f_classif), to determine the correlation between each feature and the criterion variable. The technique ranks features by their statistical significance to the difference between normal and anomalous classes.

Each dataset contains a fixed quantity of the best features (k = 50) chosen so that there is a tradeoff between computational efficiency and model performance. The less input features the model will be less prone to overfitting and be capable of generalizing to unseen data. In addition, the feature

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

selection saves a lot of training time and computation and is vital in large-scale cybersecurity datasets.

The chosen characteristics of the two datasets, UNSW-NB15 and CICIDS-2017 are converted into new feature matrices that are the input in the following modeling processes. The transformation will make sure that only informative features are saved and that irrelevant and redundant data are removed.

In general, selecting features improves the explain ability of models, the classification accuracy, and scalability of the proposed multi-level ensemble framework.

E. **Data Splitting and Scaling**

The processed datasets have been classified into training and testing datasets to analyze the performance of the models accurately. A typical train-test split method is used, with 80 per cent of the data used in training and 20 per cent in testing. This makes sure that the model will be tested on unknown data and give a real assessment of its generalization ability. Both datasets are split individually in the process and a fixed random state ensures reproducibility.

After data splitting, we use the Standard Scaler technique to scale the features. Scaling is necessary since the features of cybersecurity datasets may differ in their ranges and units and this may adversely affect model performance. Standard Scaler makes the data of a normal distribution with a mean of zero and a standard deviation of one.

Each dataset is equipped with separate scalers to maintain dataset-specific properties. The scaler is only applied to the training data and afterwards to the training and testing sets in order to prevent data leakage. This will ensure that the model will not acquire prior knowledge on the distribution of the test data.

This step will get the data ready to be fed into deep learning models through: ensuring numerical stability, faster convergence, and better performance in the process of training.

F. **Model Implementation**

The proposed anomaly detection system is a hybrid deep learning architecture comprising of three main models, one of them being a dense neural network using transfer learning, one being a Long Short-Term Memory (LSTM) network and a third being an auto-encoder. These models are aimed at representing various features of cybersecurity data with various learning mechanisms.

a) **Dense Neural Network with Transfer Learning**

The former is a transfer learning Dense Neural Network. This model builds upon the existing representations that have been trained to optimize the learning and extraction of features. The processed input data is forced through fully connection layers and the result is obtained in the form of probabilities. These probabilities are then converted into binary predictions with a threshold value set to 0.5 which allows the probabilities to be classified into normal and abnormal cases.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

b) **Long Short-Term Memory (LSTM) Model**

The second model is founded on the Long Short-Term Memory (LSTM) architecture that is developed to represent the temporal dependencies and the sequential patterns in the network traffic data. The input data is restructured into an appropriate three-dimensional format needed to model sequences. This sequential data is then fed into the LSTM model and the outputs are predicted values, which are again transformed into binary values by setting a threshold of 0.5.

c) **Auto encoder Model**

The third model is an auto encoder that is a mechanism of unsupervised detection of anomalies. It is only trained with normal instances of data to learn its underlying representation. In the testing stage, the model is rebuilt on the input data, and reconstruction error is calculated. The data points are categorized as normal and abnormal based on a threshold that is determined by the 95 th percentile of reconstruction error.

d) **Weighted Ensemble Model (Final Model)**

A fourth model is based on a weighted ensemble model, which is a combination of predictions of the three other single models. All models are weighted with the final prediction being 0.5 with dense model, 0.3 with LSTM model, and 0.2 with auto encoder. The products of the weighted outputs are summed up and transformed into final binary outputs with the help of a thresholding mechanism. This ensemble approach combines the results of each model to form a single decisionmaking system.

TABLE II. HYPERPARAMETER DETAILS

Model	Hyperparameter	Value / Setting	Description
Dense Neural Network (Transfer Learning)	Epochs	10	Number of training iterations
	Batch Size	64	Number of samples per batch
	Validation Split	0.2	20% of training data used for validation
	Activation Function	Sigmoid (Output)	Binary classification output
	Threshold	0.5	Decision boundary for classification
	Input Features	50	Selected features after feature selection
	Transfer Learning	Enabled	Uses pre-trained feature representations
	Optimizer	Adam (default)	Optimization algorithm

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

	Loss Function	Binary Crossentropy	Loss for binary classification
LSTM Model	Epochs	5	Number of training iterations
	Batch Size	64	Number of samples per batch
	Validation Split	0.2	20% validation data
	Input Shape	3D (samples, timesteps, features)	Required for sequence modeling
	Activation Function	Sigmoid (Output)	Binary classification output
	Threshold	0.5	Classification threshold
	Optimizer	Adam (default)	Optimization method
	Loss Function	Binary Crossentropy	Loss function
Autoencoder Model	Epochs	5	Training iterations
	Batch Size	64	Batch size
	Validation Split	0.2	Validation data split
	Training Data	Normal Class Only	Learns normal behavior
	Reconstruction Error	MSE	Mean Squared Error used
	Threshold	95th Percentile	Cutoff for anomaly detection
	Optimizer	Adam (default)	Optimization algorithm
	Loss Function	Mean Squared Error	Reconstruction loss
Ensemble Model (Final)	Method	Weighted Averaging	Combines predictions

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

	Dense Weight	0.5	Contribution of Dense model
	LSTM Weight	0.3	Contribution of LSTM model
	Autoencoder Weight	0.2	Contribution of Autoencoder
	Final Threshold	0.5	Final classification decision

IV. Result & Discussion

This section provides the framework of evaluation that would be employed to evaluate the performance of the proposed multi-level ensemble and transfer learning-based anomaly detection model. The analysis is divided into two major sections: the general model performance and the specific classification analysis. The standard performance measures are used to judge the model: accuracy, precision, recall, and F1-score, commonly applied in cybersecurity in the context of the detection of anomalies.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \quad (4)$$

A. Ensemble Model Performance Table III. Final Ensemble Model Accuracy

Dataset	Accuracy
UNSW-NB15	0.8268
CICIDS-2017	0.9629

The ensemble model has good performance on both datasets and is more accurate on CICIDS2017 (96.29) over UNSW-NB15 (82.68). This implies higher separability of patterns in CIC data. The findings show that the suggested weighted ensemble method is an effective generalization tool across various cybersecurity data sets that have different complexity and attack nature.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

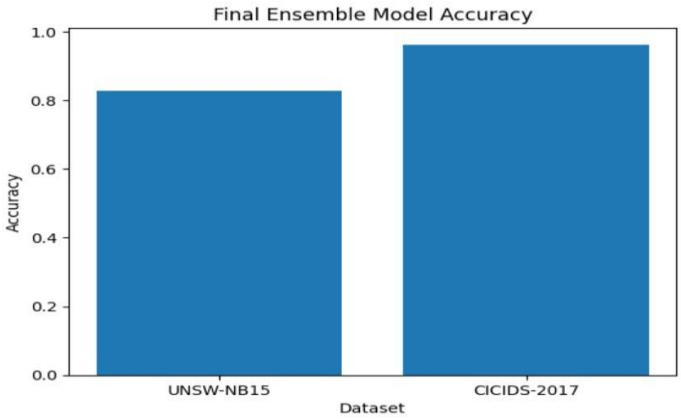

Fig. 5 Performance Graph

B. **Detailed Performance Analysis Table IV. Classification Performance – UNSW Dataset**

Class	**Precision**	**Recall**	**F1-Score**	**Support**
Normal (0)	0.76	0.95	0.84	9922
Attack (1)	0.94	0.70	0.80	10078
Overall Accuracy	**0.83**			20000
Macro Avg	0.85	0.83	0.82	20000
Weighted Avg	0.85	0.83	0.82	20000

The UNSW findings indicate that attack detection has a high level of accuracy (0.94), which means that the identified attacks are accurately predicted. But the lower recall (0.70) indicates that there are attacks that are not detected. The model is effective in detecting normal traffic (recall 0.95). In general, the outcomes indicate dataset complexity and feature a moderate degree of detection.

Table V. Classification Performance – CIC Dataset

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.94	0.99	0.96	9930
Attack (1)	0.99	0.93	0.96	10069

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

Overall Accuracy	**0.96**			19999
Macro Avg	0.96	0.96	0.96	19999
Weighted Avg	0.96	0.96	0.96	19999

The results of the CIC dataset indicate a balanced and high performance in all the metrics, with the precision and recall over 0.93 on both classes. The model can be useful in identifying both normal and attack traffic with minimal errors. The overall F1-score (0.96) shows that the model can classify well and detect anomalies with high accuracy.

C. Discussion

The experimental outcomes show that the suggested multi-level ensemble model is efficient on both datasets, and the results are much better on the CICIDS-2017 dataset than on the UNSWNB15 dataset.

In the case of the UNSW dataset, the model had an accuracy of 82.68% which implies that it has a high potential in separating normal and anomalous traffic. The accuracy of detecting an attack is very high (0.94), meaning that the majority of the attacks predicted are correctly detected. The recall of the attack class (0.70) is relatively low, though, which indicates that not all attack instances are identified. This shows how complex and variable attack patterns are in UNSW dataset. The recall of normal traffic was high (0.95) which means that the model works well with the identification of benign activity.

Conversely, CIC dataset is performing better with an accuracy of 96.29. The precision and recall rates are also high (over 0.93) in both classes, which leads to a balanced F1-score of 0.96. This shows that the model is effective in identifying normal and anomalous cases with minimum misclassification. The enhanced performance could be explained by more easily discernible feature patterns and separability in the CIC dataset.

The weighted ensemble method is an important part of these results because it incorporates the advantages of Dense (transfer learning), LSTM, and Autoencoder models. The dense model helps to make correct classification, LSTM learns sequential dependencies, and autoencoder finds anomalies based on the reconstruction error. The weights of the assigned models (0.5, 0.3, 0.2) provide even contribution of the models.

The last anomaly detection model is created through a combination of three complementary models: a Dense Neural Network with transfer learning, an LSTM network, and an autoencoder. The model produces binary predictions, which are the learning mechanism-classification probabilities of Dense and LSTM and reconstruction error of the autoencoder. A weighted ensemble method is used to combine these outputs with weights of 0.5, 0.3 and 0.2 respectively.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open access journal
Impact Factor: 8.3 www.ijesh.com ISSN: 2250 3552

The aggregate prediction is thresholded to get the final decision. This grouping model takes advantage of varying learning patterns and yields better performance with 82.68% accuracy on UNSW and 96.29% accuracy on CIC datasets, the proposed framework has a high generalization ability and applicability in various cybersecurity datasets, thus can be used in practice through anomaly detection.

TABLE VI. COMPARATIVE ANALYSIS BETWEEN EXISTING AND PROPOSED WORK

Model	Accuracy	Ref.
AD-MLA	95.3	[19]
Proposed Ensemble Model of CICIDS 2017	96.2	--

The comparative analysis shows that the proposed ensemble model is more accurate compared to the existing AD-MLA approach. The AD-MLA model has an accuracy of 95.3% whereas the ensemble model proposed has a higher accuracy of 96.2% on the CICIDS 2017 data. This enhancement shows that the combination of various learning models, such as Dense (transfer learning), LSTM, and autoencoder, into one structure is effective. The improvement in performance can be explained by the fact that the model has the capacity to capture different data patterns and minimise the model shortcomings. In general, the suggested solution is stronger and more credible in detecting anomalies in cybersecurity systems.

V. Conclusion

The research proposal offers a strong and scalable framework to detect anomalies in cybersecurity by combining data-driven preprocessing methods with state-of-the-art deep learning models. The methodology makes effective use of benchmark datasets, i.e., UNSW-NB15 and CICIDS-2017, and provides diversity and generalization to anomaly detection in networks. Comprehensive preprocessing, such as data cleaning, normalization and feature selection, are used to improve data quality and dimensionality reduction, which is used to increase model efficiency. The Ensemble Model architecture, a Dense Neural Network with transfer learning, LSTM and autoencoder, manages to capture both spatial and time patterns, and reconstruction-based anomalies. The weighted ensemble method further enriches the model integrating the strengths of the individual models and thus results in the enhancement of the detection performance and strength. Empirical results show that the model works well on datasets, and better on CICIDS-2017 because of the higher separability of the features, and still offers reliable results on the more challenging UNSW dataset. The comparative analysis proves that the proposed model is better than current methods as it is more accurate and can be more generalized. On the whole, the framework can be seen as a

reliable, efficient, and versatile solution to real-world cybersecurity anomaly detection, and in the future, it could be improved to work in real-time and deal with emerging cyber threats.

References

[1] D. Abshari and M. Sridhar, "A Survey of Anomaly Detection in Cyber-Physical Systems," 2025, [Online]. Available: http://arxiv.org/abs/2502.13256

[2] H. Huang, P. Wang, J. Pei, J. Wang, S. Alexanian, and D. Niyato, "Deep Learning Advancements in Anomaly Detection: A Comprehensive Survey," *IEEE Internet Things J.*, vol. 12, no. 21, pp. 44318–44342, 2025, doi: 10.1109/JIOT.2025.3585884.

[3] G. Andronikidis *et al.*, "AI-Driven Anomaly and Intrusion Detection in Energy Systems: Current Trends and Future Direction," *Proc. 2024 IEEE Int. Conf. Cyber Secur. Resilience, CSR 2024*, pp. 777–782, 2024, doi: 10.1109/CSR61664.2024.10679380.

[4] Q. U. Ain, A. Aftab Ahmed Jilani, N. Azhar Butt, S. Ur Rehman, and H. Abdulrahman Alhulayyil, "Anomaly Detection for Aviation Cyber-Physical System: Opportunities and Challenges," *IEEE Access*, vol. 12, no. November, pp. 175905–175925, 2024, doi: 10.1109/ACCESS.2024.3495519.

[5] K. H. Tang, M. C. Ghanem, P. Gasiorowski, V. Vassilev, and K. Ouazzane, "Synchronisation, Optimisation and Adaptation of Machine Learning Techniques for Computer Vision in Cyber-Physical Systems: A Comprehensive Analysis," *IET Cyber-Physical Syst. Theory Appl.*, vol. 10, no. 1, 2025, doi: 10.1049/cps2.70031.

[6] P. Sajjadi, F. Dinmohammadi, and M. Shafiee, "Machine Learning in Prognostics and Health Management of Cyber-Physical Systems: A Review," *IEEE Access*, vol. 13, no. August, pp. 162320–162354, 2025, doi: 10.1109/ACCESS.2025.3605553.

[7] C. M. Ahmed, M. R. Gauthama Raman, and A. P. Mathur, "Challenges in Machine Learning based approaches for Real-Time Anomaly Detection in Industrial Control Systems," *CPSS 2020 - Proc. 6th ACM Cyber-Physical Syst. Secur. Work. Co-located with AsiaCCS 2020*, no. October 2020, pp. 23–29, 2020, doi: 10.1145/3384941.3409588.

[8] S. Jamshidi, A. Amirnia, A. Nikanjam, and F. Khomh, "Enhancing Security and Energy Efficiency of Cyber-Physical Systems using Deep Reinforcement Learning," *Procedia Comput. Sci.*, vol. 238, no. 2019, pp. 1074–1079, 2024, doi: 10.1016/j.procs.2024.06.137.

[9] S. Wang, R. K. L. Ko, G. Bai, N. Dong, T. Choi, and Y. Zhang, "Evasion Attack and Defense on Machine Learning Models in Cyber-Physical Systems: A Survey," *IEEE Commun. Surv. Tutorials*, vol. 26, no. 2, pp. 930–966, 2024, doi: 10.1109/COMST.2023.3344808.

[10] S. K. Hussein and M. A. El-Dosuky, "Anomaly Detection in Cyber-Physical Systems Using Explainable Artificial Intelligence and Machine Learning," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 8, pp. 3138–3151, 2023.

[11] S. Guarino, F. Vitale, F. Flammini, L. Faramondi, N. Mazzocca, and R. Setola, "A TwoLevel Fusion Framework for Cyber-Physical Anomaly Detection," *IEEE Trans. Ind. CyberPhysical Syst.*, vol. 2, pp. 1–13, 2023, doi: 10.1109/ticps.2023.3336608.

[12] M. Archana, S. Kavitha, and A. Vani Vathsala, "Auto Deep Learning-based Automated Surveillance Technique to Recognize the Activities in the Cyber-Physical System," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 2, pp. 35–42, 2023, doi: 10.17762/ijritcc.v11i2.6111.

[13] Y. Luo, Y. Xiao, L. Cheng, G. Peng, and D. D. Yao, "Deep Learning-based Anomaly Detection in Cyber-physical Systems: Progress and Opportunities," *ACM Comput. Surv.*, vol. 54, no. 5, 2022, doi: 10.1145/3453155.

[14] S. M. Nagarajan, G. G. Deverajan, A. K. Bashir, R. P. Mahapatra, and M. S. Al-Numay, "IADF-CPS: Intelligent Anomaly Detection Framework towards Cyber Physical Systems," *Comput. Commun.*, vol. 188, pp. 81–89, 2022, doi: 10.1016/j.comcom.2022.02.022.

[15] C. Feng and P. Tian, "Time Series Anomaly Detection for Cyber-physical Systems via Neural System Identification and Bayesian Filtering," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, no. August 2021, pp. 2858–2867, 2021, doi: 10.1145/3447548.3467137.

[16] R. Aziz *et al.*, "Critical Evaluation of Data Privacy and Security Threats: An Intelligent Federated Learning-based Intrusion Detection System Poisoning Attack and Defense for Cyber- Physical Systems its Issues and Challenges Related to Privacy and Security in IoT," *Asian Bull. Big Data Manag.*, vol. 5, no. 1, pp. 73–84, 2025, doi: 10.62019/abbdm.v5i1.293.

[17] G. Ahmadi-Assalemi, H. Al-Khateeb, V. Benson, B. Adamyk, and M. Ammi, "Adaptive learning anomaly detection and classification model for cyber and physical threats in industrial control systems," *IET Cyber-Physical Syst. Theory Appl.*, vol. 10, no. 1, pp. 1–17, 2025, doi: 10.1049/cps2.70004.

[18] M. A. S. P. Dayarathne, M. S. M. Jayathilaka, R. M. V. A. Bandara, V. Logeeshan, S. Kumarawadu, and C. Wanigasekara, "Mitigating Cyber Risks in Smart Cyber-Physical Power Systems Through Deep Learning and Hybrid Security Models," *IEEE Access*, vol. 13, no. March, pp. 37474–37492, 2025, doi: 10.1109/ACCESS.2025.3545637.

[19] W. K. Wong, S. Baskar, K. M. Abubeker, and P. K. Ng, "Sustainable cyber-physical VANETs with AI-driven anomaly detection and energy-efficient multi-criteria routing

using machine learning algorithms," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-28212-1. [20] M. D. Sikder, "AI Methods for Anomaly Detection in Cyber-Physical Systems: With Application to Water and Agriculture," 2025, [Online]. Available: https://vtechworks.lib.vt.edu/items/ee8fb850-a0e7-4b54-ad17-0b0ef8d43d40

[21] H. S. Mavikumbure *et al.*, "Cy-Phy ADS: Cyber Physical Anomaly Detection Framework for EV Charging Systems," *IEEE Trans. Transp. Electrif.*, vol. PP, no. January, p. 1, 2024.

[22] S. Gaba *et al.*, "A Systematic Analysis of Enhancing Cyber Security Using Deep Learning for Cyber Physical Systems," *IEEE Access*, vol. 12, no. December 2023, pp. 6017–6035.

[23] N. Nabi, M. M. Rahman, S. K. Ghosh, S. Alam, and R. Barua, "Machine Learning-Based Anomaly Detection for Cyber Threat Prevention," *J. Primeasia*, vol. 6, no. 1, pp. 1–8, 2024.

[24] B. Bajic, A. Rikalovic, N. Suzic, and V. Piuri, "Toward a Human-Cyber-Physical System for Real-Time Anomaly Detection," *IEEE Syst. J.*, vol. 18, no. 2, pp. 1308–1319, 2024.

[25] D. M. Manias *et al.*, "Trends in Smart Grid Cyber-Physical Security: Components, Threats, and Solutions," *IEEE Access*, vol. 12, no. November, pp. 161329–161356, 2024.