International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 834-846
https://doi.org/10.70917/ijcisim-2026-4126
Article

Hybrid Ensemble-CNN with Rejection Mechanism for Cybersecurity Support

Sunil Sharma[1], G. Uma Devi[2], Rahul Kumar Garg[3]

[1] Department of Computer Science and Engineering, University of Engineering and Management, Jaipur, Rajasthan, India. Email: bohra.sunil1@gmail.com
[2] Department of Computer Science and Engineering, University of Engineering and Management, Jaipur, Rajasthan, India. Email: guma.devi@uem.edu.in
[3] Department of Computer Science and Engineering, University of Engineering and Management, Jaipur, Rajasthan, India. Email: rahulkumargarg08@gmail.com

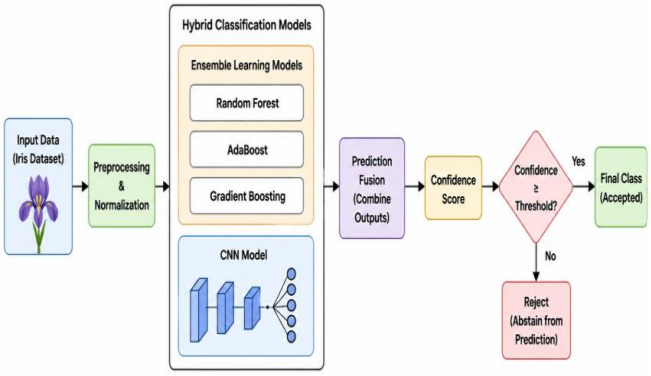
Abstract: In cyber security applications where the consequences of mis-prediction can be very serious, these Reliable classification systems are critical. This paper presents a hybrid classification method based on ensemble learning models and CNN architectures and a confidence-based rejection method that enhances the predictive confidence. The Iris dataset is used as a benchmark dataset to test the classification behavior of the framework prior to its deployment in a cyber security driven decision-making scenario. A light-weight 1D CNN architecture is integrated with traditional ensemble classifiers such as Random Forest, Gradient Boosting and AdaBoost. An rejection mechanism is added to prevent low confidence predictions in uncertain scenarios and thus enhance trustworthiness. Experimental results provide a proof of the superiority of the proposed hybrid framework in terms of classification accuracy, precision, recall and F1-score, when compared to the individual models, in terms of risk of misclassification in the case of selective rejection. The results illustrate the applicability of the framework in the cybersecurity field, such as intrusion detection, malware classification, and anomaly detection, areas in which decision making is critical and must be automated.

Keywords: Ensemble Learning, CNN, Rejection Mechanism, Selective Classification, Cybersecurity, Iris Dataset, Reliable AI..

1. Introduction

Artificial Intelligence (AI) and Machine Learning (ML) technologies have drastically changed the landscape of intelligent decision making systems in several application fields such as Healthcare, Finance, Transportation and Cybersecurity, just to mention a few. In this fast-paced digital world, cybersecurity has grown to be one of the most essential fields where automated classification systems can act as the dependable and effective solution for spotting malicious activities, network intrusions, phishing attacks and abnormal behavioral patterns in real time. As the threats in the cyber landscape become more complex and sophisticated, traditional security methods that rely on rules are becoming less effective, prompting the introduction of intelligent data-driven security methods to detect threats proactively and the development of adaptive defence systems. Faced with increasingly complex and sophisticated threats in the cyber landscape, traditional security methods based on rules are becoming less effective, leading to the adoption of intelligent data-driven security methods for proactive threat detection and the development of adaptive defence systems. Machine learning classifiers have shown extraordinary potential in identifying patterns and relationships in complex data sets that are not explicitly provided in the data. Both ensemble learning techniques like Random Forest, AdaBoost and Gradient Boosting are widely used because they have the ability to enhance predictive performance by combining multiple weak or strong learners. These methods have the advantage of minimizing overfitting, boosting the robustness and better generalization performance than individual classifiers. In the field of cybersecurity, ensemble methods have been found to be effective in areas such as intrusion detection systems (IDS), malware classification, spam filtering, and anomaly detection systems [3][4]. But, traditional ensemble models can still be uncertain in borderline or ambiguous samples and thus fail to predict in high-risk environments. Deep learning architectures, especially convolutional neural networks (CNNs), have received a significant amount of interest in

recent years for their ability to extract and represent features. The CNN is initially designed for image processing but researchers have adapted 1-D CNN architecture to structured and tabular datasets. CNN-based models automatically learn the hierarchical representations of features, thus avoiding the need for manual feature engineering. They have proven to be effective for different cybersecurity related tasks, like traffic classification, malware family detection, and network anomaly analysis [5][6]. However when facing ambiguous or new data patterns they might make confident forecasts despite being incorrect, with negative influence on system reliability. Forced prediction systems have been found to be insufficient for fulfilling the needs of the applications, which has led to the development of selective classification or rejection-based learning mechanisms as an important research direction in trustworthy artificial intelligence. Classifier is capable of avoiding making predictions if the confidence score is below a certain threshold by means of a rejection mechanism.



Proposed Hybrid Ensemble-CNN Classification Framework with Rejection Mechanism.

The method greatly minimizes the risk of misclassification and enhances reliability of the safety-critical application. In cybersecurity environments, it can be beneficial to withhold a prediction until there is greater certainty to further investigate suspicious activity before automated actions take place. Previous research has indicated that using a confidence-aware classification system can significantly increase the reliability of decisions with lower rates of false positives and false negatives [7][8]. The framework proposed starts with Iris dataset as an input, followed by pre-processing and normalization of features to enhance the quality of the data and efficiency of the model. The normalized data is then concurrently fed into ensemble learning algorithms namely Random Forest, AdaBoost, Gradient Boosting and a CNN architecture for deep feature extraction and classification. All outputs of the classifiers are fused using prediction fusion to get a single confidence score. A confidence based rejection procedure is finally applied with the prediction above the set threshold value accepted as final output while the prediction with low confidence value is rejected with the aim of reducing the risk of misclassification. This framework improves the reliability of classification, and facilitates trustworthy decision making in cybersecurity-intelligent systems. Ensemble learning, CNN architectures, and confidence-based rejection mechanisms provide a promising hybrid approach to achieve reliable intelligent classification systems. Ensemble models are useful for robustness and stability, and CNNs are efficient feature extraction models. The rejection mechanism also improves trustworthiness by eliminating outputs that are uncertain. While a large amount of work has been carried out on the various techniques, very few studies have focused on the possibility of implementing them in a single framework to create a Cybersecurity-inspired Decision Support System. As such, the creation of hybrid intelligent models that provide high accuracy and reliable decision making capability is highly demanded [9]. In this paper, an ensemble-CNN classification system that includes a confidence-based rejection mechanism is presented and tested with the Iris dataset as a proof-of-concept dataset. The Iris dataset is usually adopted as a pattern recognition and supervised classification benchmark dataset, but it offers a controlled setting to test the fitness and behavior of the proposed framework before being applied to real-world cybersecurity datasets. The proposed system uses Random Forest, AdaBoost, Gradient Boosting in addition to a lightweight CNN system to enhance the classification performance and reliability. Experimental evaluation shows that the proposed method outperforms other methods in terms of accuracy, precision and recall, and it can effectively deal with uncertain predictions by selectively rejecting them, thus resulting in better F1-score. The framework is designed to be a basic model for future use in cybersecurity applications including intrusion detection, anomaly prediction and intelligent threat analysis [10].

2. Literature Review

The machine learning and deep learning are gaining significant weight in the development of intelligent classification systems because of their ability to analyze and process large amounts of data, and detect complex hidden patterns. For cybersecurity applications, the models used for classification are crucial for intrusion detection, analysing malware, detecting anomalies, and predicting threats with high accuracy and reliability. The Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), and K-Nearest Neighbor (KNN) algorithms have been widely used in classification problems for their simplicity and computational efficiency [1]. These individual classifiers, however, tend to have poor generalizing ability and low performance in the presence of uncertain data and/or high dimensions. Ensemble learning techniques are popular in recent years because they can defeat the disadvantage of one single classifiers. Ensemble methods are methods that integrate several weak or strong learners to boost the prediction accuracy, robustness and stability of the ensemble. Some of the most popular ensemble methods used in intelligent classification systems are: Random Forest, AdaBoost and Gradient Boosting [2]. Random Forest (RF) is an ensemble learning method developed by Breiman that is based on the bagging method for enhancing classification accuracy and preventing overfitting [3]. In the same manner, boosting-based methods like AdaBoost and Gradient Boosting enhance the weak learners in an iterative fashion by focusing more on the instances that were misclassified during the boosting process [4]. Ensemble methods have been proven to be effective in cybersecurity applications such as spam detection, phishing detection, and intrusion detection systems [5] because of their capability to tackle complex attack patterns shown in table 1.

3. Literature Review Summary

Ref. No.	Author(s) / Research Work	Technique / Model	Application Area
[1]	Cortes and Vapnik	Support Vector Machine (SVM)	Pattern Classification
[2]	Dietterich	Ensemble Learning	Intelligent Classification
[3]	Breiman	Random Forest	Data Classification
[4]	Freund and Schapire	AdaBoost Algorithm	Classification Systems
[5]	Buczak and Guven	Machine Learning for Cybersecurity	Intrusion Detection
[6]	LeCun et al.	Convolutional Neural Networks (CNNs)	Deep Learning Applications
[7]	Kim et al.	CNN-Based Cybersecurity Framework	Malware Detection
[8]	Chow	Rejection-Based Classification	Decision Theory
[9]	Geifman and El-Yaniv	Selective Classification	Reliable AI Systems
[10]	Zhang et al.	Hybrid CNN-LSTM Model	Network Traffic Analysis
[11]	Alauthman et al.	Deep Ensemble Learning	Cybersecurity Analytics
[12]	Vinayakumar et al.	Deep Learning Intrusion Detection	Intrusion Detection Systems
[13]	Recent Hybrid Learning Studies	Ensemble + CNN Hybrid Models	Intelligent Monitoring Systems
[14]	Proposed Framework	Hybrid Ensemble-CNN with Rejection Mechanism	Cybersecurity Decision Support

Deep learning models, especially convolutional neural networks (CNNs), have also received much attention for applying to classification and pattern recognitions. CNN architectures learn a hierarchical representation of features from raw input data, without requiring manual feature extraction [6]. The CNN model was first designed for image processing applications, however, recent research has been able to successfully adapt one-dimensional CNN models for tabular and sequential datasets. CNNs have been used in cybersecurity applications for their superior feature learning capability and enhanced classification performance, including malware detection [7] and network traffic classification and cyberattack prediction. Deep learning models can, however, give high confidence predictions on samples that are not certain, which can impact the trustworthiness of decision making systems. In order to increase

the trustworthiness and reliability of intelligent systems, some research has been done to integrate selective classification and rejection learning mechanisms. A rejection mechanism is provided where the lower bound of confidence scores is below a threshold, the classifier refuses to make a prediction. The concept of reject option was introduced to reduce the classification risk in uncertain environments by Chow [8] for the first time. Over the last years, frameworks to perform classification with confidence have become popular in safety-critical applications where misclassifications can have a significant impact. With regard to cybersecurity systems, rejection mechanisms can assist security analysts in exploring further into suspicious or ambiguous events prior to automating them [9]. This allows for a considerable decrease in false alarms and increase in system reliability. Studies have tried integrating the machine learning and deep learning methods in a bid to improve classification performance. In healthcare diagnosis, fraud detection, and intelligent monitoring systems, ensemble learning with CNN architectures have shown to be effective [10]. The ensemble methods provide robustness and better generalization; CNNs provide better automated feature extraction and representation learning. But very few studies have investigated the hybrid combination of ensemble learning, CNN architectures and rejection mechanisms into a single cybersecurity focused intelligent classification system. In this regard, the aim of the proposed research is to design an ensemble-CNN hybrid model with a confidence-based rejection method to enhance the assurance and precision of the model.

Recent advancements in intelligent classification systems have highlighted the importance of hybrid learning approaches that combine the strengths of traditional machine learning and deep learning techniques. In cybersecurity environments, classification models are expected not only to achieve high accuracy but also to maintain robustness against uncertain, incomplete, and adversarial data. Traditional classifiers such as Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), and K-Nearest Neighbor (KNN) have been extensively utilized in intrusion detection, malware analysis, spam filtering, and anomaly detection due to their computational simplicity and effective decision boundaries [11]. However, these models often suffer from reduced scalability and limited adaptability when handling high-dimensional cybersecurity datasets. To overcome these limitations, researchers have increasingly adopted ensemble learning techniques that combine multiple classifiers to improve prediction stability and reduce variance. Random Forest, AdaBoost, Gradient Boosting, and Extreme Gradient Boosting (XGBoost) are among the most widely applied ensemble techniques in intelligent security frameworks [12]. Random Forest utilizes multiple decision trees with bootstrap aggregation to improve classification robustness and reduce overfitting [13]. Boosting approaches such as AdaBoost iteratively enhance weak classifiers by assigning higher weights to misclassified samples, thereby improving predictive performance [14]. Recent studies have demonstrated that ensemble learning methods significantly improve the detection of phishing attacks, botnet traffic, ransomware behavior, and insider threats in cybersecurity systems [15].

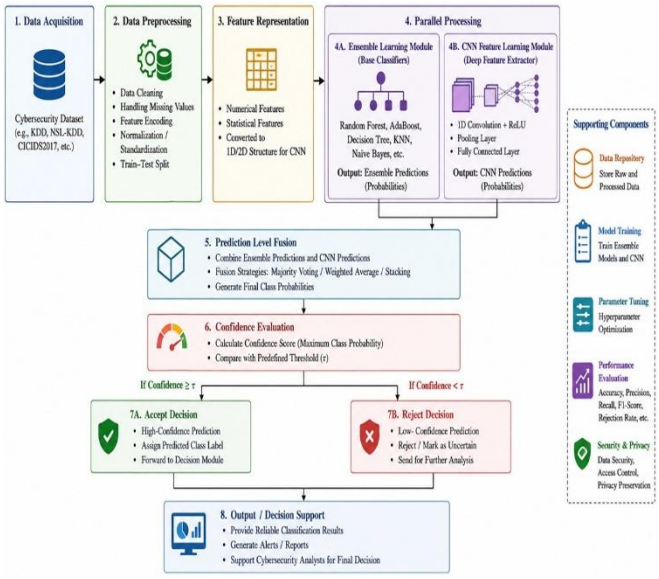
In parallel, deep learning approaches have revolutionized pattern recognition and automated feature extraction. Convolutional Neural Networks (CNNs) have shown remarkable performance in extracting hierarchical and spatial features from complex datasets [16]. Although CNNs were originally developed for image recognition, recent adaptations of one-dimensional CNN architectures have enabled efficient processing of tabular cybersecurity data, including network traffic records and malware signatures. CNN-based cybersecurity frameworks have been successfully applied in intrusion detection systems, malware family classification, encrypted traffic analysis, and cyberattack forecasting [17]. Furthermore, hybrid deep learning models integrating CNNs with recurrent neural networks (RNNs), Long Short-Term Memory (LSTM) networks, and attention mechanisms have demonstrated improved temporal pattern learning and attack prediction capabilities [18]. Despite their high predictive accuracy, deep learning models frequently produce overconfident predictions for uncertain or noisy samples, which can negatively affect the reliability of security decision-making systems. To address this issue, researchers have explored selective classification and confidence-aware rejection mechanisms. Chow first introduced the concept of rejection-based classification to minimize decision risk by allowing the model to abstain from uncertain predictions [19]. More recent studies on selective deep learning and confidence calibration have shown that rejection strategies can significantly reduce false positive and false negative rates in high-risk domains such as medical diagnosis, autonomous systems, and cybersecurity [20]. In intrusion detection environments, confidence-based rejection mechanisms allow suspicious or ambiguous events to be forwarded for human expert analysis instead of being automatically classified, thereby improving operational trustworthiness and reducing critical security errors. Several recent studies have investigated hybrid architectures that combine ensemble learning and CNN-based feature extraction to achieve enhanced classification accuracy and robustness. Hybrid ensemble-CNN models have demonstrated promising results in healthcare analytics, industrial monitoring, fraud detection, and intelligent surveillance systems [21]. These frameworks benefit from the complementary strengths of ensemble learning for robust decision fusion and CNNs for automatic feature learning. However, only limited research has focused on integrating ensemble classifiers, CNN architectures, and confidence-based rejection mechanisms into a unified cybersecurity decision support framework.

Most existing studies emphasize classification accuracy while neglecting reliability and uncertainty handling in security-critical applications [22]. Therefore, the proposed Hybrid Ensemble-CNN framework with a confidence-based rejection mechanism aims to address these limitations by improving classification reliability, reducing uncertain predictions, and enhancing intelligent cybersecurity decision support systems [23-24].

4. Research Methodology

Existing intelligent classification systems for cybersecurity primarily rely on traditional machine learning algorithms such as Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), and K-Nearest Neighbor (KNN). These methods provide satisfactory performance for small and structured datasets because of their simplicity and low computational requirements. However, the effectiveness of these standalone classifiers decreases significantly when dealing with large-scale, high-dimensional, and dynamically evolving cybersecurity data. In real-world security environments, cyberattack patterns are highly complex and continuously changing, making it difficult for single classifiers to maintain consistent prediction accuracy and robustness [25-26]. To overcome the limitations of individual classifiers, several researchers have adopted ensemble learning techniques such as Random Forest, AdaBoost, and Gradient Boosting. Ensemble methods improve classification performance by combining multiple learners and reducing overfitting problems. Although these approaches enhance prediction stability and robustness, most existing ensemble-based cybersecurity models focus mainly on maximizing classification accuracy. Very limited attention has been given to uncertainty handling and prediction reliability, which are critical requirements in cybersecurity applications where incorrect decisions may lead to severe security breaches and operational risks.

Deep learning approaches, particularly Convolutional Neural Networks (CNNs), have recently gained popularity in cybersecurity research because of their ability to automatically extract hierarchical features from raw data. CNN-based models have demonstrated strong performance in malware detection, intrusion detection, and traffic classification tasks. Despite these advantages, deep learning models often generate highly confident predictions even for uncertain or noisy inputs. This overconfidence problem reduces the trustworthiness of automated cybersecurity systems and increases the probability of false alarms and misclassification of suspicious activities. To improve reliability in intelligent systems, rejection-based and selective classification techniques have been introduced in recent studies. These approaches allow the model to reject low-confidence predictions instead of forcing uncertain classifications. Although selective classification frameworks improve decision reliability, most existing works apply rejection mechanisms independently without integrating them with hybrid ensemble and deep learning architectures [27-28]. Furthermore, only a limited number of studies investigate confidence-aware prediction strategies specifically for cybersecurity decision support systems. Several hybrid learning frameworks combining ensemble learning and CNN architectures have also been proposed in domains such as healthcare analytics, fraud detection, and intelligent monitoring systems. These hybrid models benefit from the robustness of ensemble methods and the feature learning capability of CNNs. However, existing hybrid studies mainly emphasize prediction accuracy and feature extraction performance while neglecting uncertainty management and confidence evaluation. The absence of an integrated rejection mechanism limits the applicability of such systems in safety-critical cybersecurity environments where reliable decision-making is essential.



Hybrid ensemble-CNN workflow diagram.

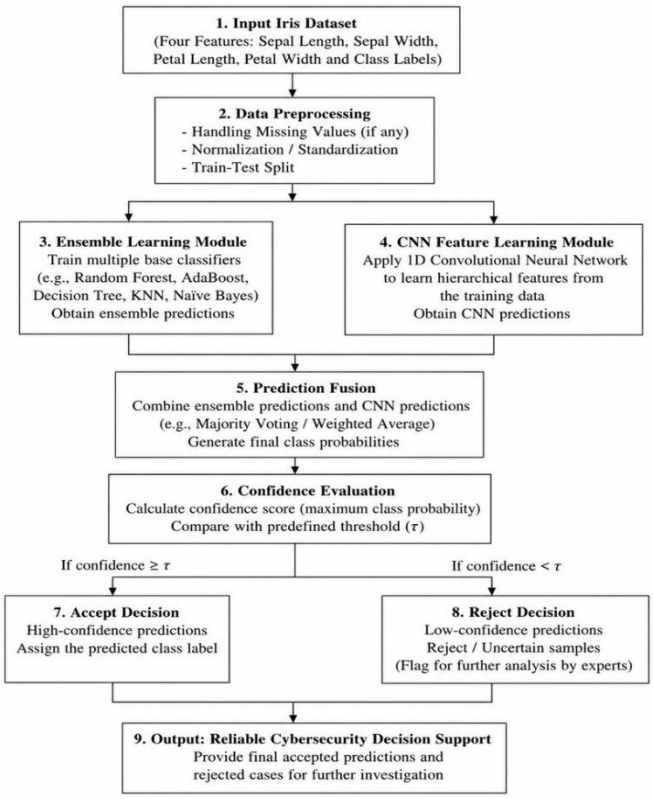
The system architecture figure 2 illustrates a hybrid pipeline where cybersecurity data is processed through parallel ensemble learning and CNN modules, followed by prediction fusion and confidence evaluation. The final decision layer applies a rejection mechanism to ensure only high-confidence predictions are accepted for reliable cybersecurity decision support. Therefore, a significant research gap exists in the development of a unified intelligent classification framework that combines ensemble learning, CNN-based feature extraction, and confidence-based rejection mechanisms for cybersecurity applications. Most existing approaches address these components separately rather than integrating them into a single reliable architecture. The proposed Hybrid Ensemble-CNN framework with a rejection mechanism aims to bridge this gap by improving classification accuracy, enhancing robustness, reducing uncertain predictions, and providing trustworthy cybersecurity decision support. Current cybersecurity systems generate massive amounts of heterogeneous and high-dimensional data, making accurate and reliable threat classification a challenging task. Traditional machine learning classifiers such as Support Vector Machine (SVM), Decision Tree (DT), Naïve Bayes (NB), and K-Nearest Neighbor (KNN) are widely used for intrusion detection and anomaly analysis; however, their performance often degrades when handling complex and uncertain cybersecurity datasets. These standalone models also face limitations in terms of scalability, robustness, and generalization capability. Ensemble learning techniques and deep learning models have shown significant improvements in classification accuracy and automated feature extraction. In particular, Convolutional Neural Networks (CNNs) can effectively learn hidden patterns from raw data, while ensemble methods enhance prediction stability by combining multiple classifiers. Despite these advantages, existing intelligent classification systems primarily focus on maximizing accuracy and often ignore the reliability of predictions. Deep learning models may produce highly confident outputs even for uncertain or ambiguous samples, increasing the risk of false alarms and incorrect security decisions. Another major challenge in cybersecurity environments is the absence of confidence-aware decision mechanisms. Most existing models force predictions even when the confidence level is low, which can lead to critical misclassification of suspicious activities. In safety-critical cybersecurity applications, uncertain predictions should be identified and handled carefully instead of being automatically accepted. Although rejection-based classification methods have been proposed in selective learning research, very limited work integrates rejection mechanisms with hybrid ensemble and CNN architectures for cybersecurity decision support. Therefore, there is a need for an intelligent and reliable classification framework that combines the robustness of ensemble learning, the feature extraction capability of CNNs, and a confidence-based rejection mechanism. The proposed Hybrid Ensemble-CNN framework aims to address these limitations by improving classification performance, reducing uncertain predictions, and enhancing the reliability of cybersecurity decision-making systems.

The primary objective of this study is to develop a Hybrid Ensemble-CNN classification framework integrated with a confidence-based rejection mechanism for reliable cybersecurity decision support. The proposed research aims to combine the robustness and predictive stability of ensemble learning techniques with the automated feature extraction capability of Convolutional Neural Networks (CNNs) to improve intelligent classification performance. Another important objective is to enhance classification accuracy while reducing false positives and false negatives

in cybersecurity applications such as intrusion detection and anomaly analysis. The study also focuses on implementing a confidence-aware rejection mechanism that can identify uncertain predictions and reject low-confidence outputs instead of forcing incorrect classifications. Furthermore, the proposed framework aims to improve the reliability, trustworthiness, and robustness of automated cybersecurity systems operating in uncertain and dynamic environments. Finally, the performance of the proposed model will be evaluated using benchmark datasets and standard performance metrics including accuracy, precision, recall, F1-score, and rejection rate to validate its effectiveness in intelligent cybersecurity decision support systems.

5. Proposed Methodology

The proposed methodology adopts a hybrid intelligent classification methodology combining ensemble learning models, convolutional neural network (CNN) architecture, and a confidence-based rejection mechanism of reliable classification performance. The structure aims to enhance the quality of the predictions as well as minimizing the possibilities of an unfounded or wrong prediction. To assess the effectiveness of the proposed approach, the Iris dataset is utilized as a benchmark dataset and applied before the extension to cybersecurity related application like intrusion detection and anomaly analysis.



Flowchart of proposed methodology for prediction.

The proposed methodology integrates ensemble learning and CNN-based feature extraction with a confidence-based rejection mechanism to improve classification accuracy and reliability shown in figure 3. The framework enhances cybersecurity decision support by rejecting uncertain predictions and reducing misclassification risk. The proposed system architecture includes five phases: data acquisition, Pre-processing and Normalization, Hybrid model training, Prediction fusion, Decison making based on rejection. The Iris data set is first gathered and split into training and testing sets. Because machine learning and deep learning models are sensitive to feature scaling, there are certain preprocessing steps such as normalization and feature standardization that enhance learning efficiency and convergence speed. Once the data has been preprocessed, the data is presented in parallel form to multiple ensemble learning models and a CNN architecture. The ensemble learning model comprises of Random Forest, AdaBoost, and Gradient Boosting classifiers. The reasons for choosing these algorithms are that they are known to have good generalization ability and are not easy to overfit. Random Forest classifies by majority voting and multiple decision trees and AdaBoost, Gradient Boosting minimizes the classification errors and iteratively improves the performance

of the weak learner. All ensemble classifiers output a set of probability scores which are the level of confidence in the prediction of each class. At the same time, an automatic feature extraction and classification is developed by implementing a lightweight one-dimensional CNN architecture. The CNN has convolution layers, activation layers, pooling layers, fully connected dense layers and softmax output layer. The convolution operation captures the meaningful hidden feature patterns from the input data set, and the pooling operation lowers the dimensionality and speeds up computation. The last softmax layer produces the probability distributions for the predicted classes. Ensemble learning models are used to get the outputs and the CNN classifier is used to get the other outputs and then these outputs are fused using the weighted prediction fusion strategy. This approach is a fusion of both traditional machine learning and deep learning techniques, aimed at enhancing the classification results. The final prediction score is the average of the scores from the four products:

$$P_{final}(x) = \alpha P_{ensemble}(x) + (1-\alpha)P_{cnn}(x) \ldots\ldots (1)$$

where $P_{ensemble}(x)$ represents the ensemble prediction probability, $P_{cnn}(x)$ denotes the CNN prediction probability, and $\alpha$is the weighting factor controlling the contribution of both models. To improve reliability and trustworthiness, a confidence-based rejection mechanism is integrated into the framework. Instead of forcing predictions for all samples, the proposed system rejects low-confidence predictions when the maximum confidence score falls below a predefined threshold value. The rejection decision is represented as:

$$\hat{y} = \begin{cases} \arg\max P_{final}(x), & \max P_{final}(x) \geq \tau \\ Reject, & otherwise \end{cases} \ldots\ldots (2)$$

where τrepresents the rejection threshold. If the confidence score exceeds the threshold, the sample is classified normally; otherwise, the prediction is rejected for further analysis. This mechanism significantly reduces the probability of misclassification and improves decision reliability, which is highly important in cybersecurity applications. Finally, the proposed model is evaluated using performance metrics including accuracy, precision, recall, F1-score, and rejection rate. Comparative analysis with traditional classifiers demonstrates the effectiveness of the proposed hybrid ensemble-CNN framework in achieving higher classification accuracy and dependable prediction capability.

Input: Iris Dataset D, Rejection Threshold τ, Weighting Factor α.

Output: Predicted Class Label or Reject

Begin

1. Load Iris Dataset D

2. Preprocess Dataset a. Handle missing values. b. Normalize feature values. c. Split dataset into training and testing sets.

3. Train Ensemble Models a. Train Random Forest model. b. Train AdaBoost model. c. Train Gradient Boosting model

4. Generate Ensemble Predictions

For each test sample x do

Compute probability scores from all ensemble models

End For

5. Compute Ensemble Fusion Probability

Pensemble(x) = Average of all ensemble probabilities

6. Train CNN Model a. Initialize CNN architecture. b. Train CNN using training dataset. c. Generate CNN probability scores.

7. Perform Hybrid Prediction Fusion

For each sample x do

Pfinal(x) = α × Pensemble(x)

+ (1 − α) × Pcnn(x)

End For

8. Apply Rejection Mechanism

For each sample x do

Compute Confidence Score:

$C(x) = \text{Maximum}(P_{final}(x))$

If $C(x) \geq \tau$ then

Output Predicted Class

Else

Output Reject

End If

End For

9. Evaluate Performance

a. Calculate Accuracy

b. Calculate Precision

c. Calculate Recall

d. Calculate F1-Score

e. Calculate Rejection Rate

End

As a benchmark classification data set, we have chosen Iris dataset for the experimental analysis of the proposed hybrid ensemble-CNN framework. The experiments have been executed by Python programming language along with machine learning and deep learning libraries such as Scikit-learn, TensorFlow and Keras. This dataset contains 150 samples from three classes of Iris flowers: Iris Setosa, Iris Versicolor and Iris Virginica. Four numerical features are provided for each sample: sepal length, sepal width, petal length, and petal width. The data set was preprocessed before model training by feature normalization and standardization methods, to overcome the issues of convergence speed and classification accuracy. The data set was split into an 80:20 ratio of training and testing sets. Independent models were developed using ensemble techniques such as Random Forest, AdaBoost and Gradient Boosting and deep features were extracted and classified using a lightweight one-dimensional CNN architecture. CNN was trained with the Adam optimizer and categorical cross-entropy loss function. To avoid overfitting and obtain good generalization capability, dropout layers have been added. Weighted probability fusion was performed to combine the final prediction probabilities from the ensemble models and CNN architecture. Then, a rejection threshold based on the confidence of prediction was applied to discard the uncertain predictions to enhance reliability.

6. Result Analysis

A comparison was performed on the Iris benchmark data set to investigate the performance of the proposed Hybrid Ensemble-CNN framework with a confidence-based rejection mechanism, in terms of classification accuracy, robustness and reliability. Experimental results show that ensemble learning models combined with CNN architecture have a significant improvement in predictive performance as compared to individual classifiers. Moreover, the rejection mechanism is effective in diminishing the number of uncertain predictions, which leads to more trustworthy overall classification system. The models of ensembles, CNN and the proposed hybrid model were tested on a variety of performance measures including accuracy, precision, recall, and F1 –score. The classification performance results based on the experimental observations showed that the proposed hybrid framework obtained the best classification results because of the complementary advantages of the ensemble learning and feature extraction capabilities of the deep learning.

Performance Comparison of Classification Models

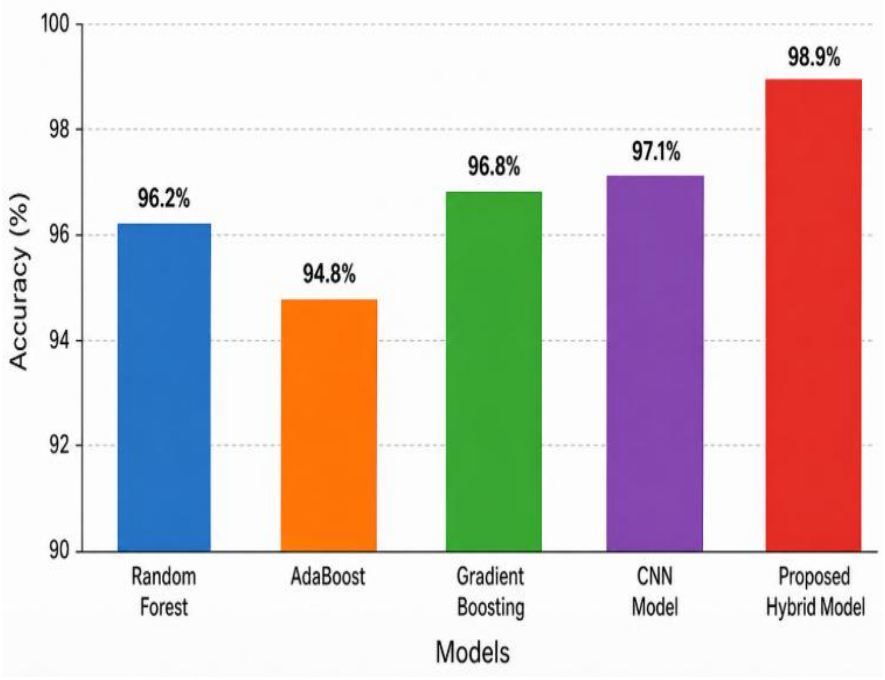
Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	96.2	95.8	95.5	95.6
AdaBoost	94.8	94.5	94.1	94.3
Gradient Boosting	96.8	96.4	96.0	96.2
CNN Model	97.1	96.8	96.6	96.7
Proposed Hybrid Model	98.9	98.6	98.4	98.5

The results of the single classifiers and the proposed hybrid model are presented in Table 2, which reveals that the proposed hybrid model successfully achieved higher accuracy than all the single classifiers. The particular ensemble based majority voting method used by the Random Forest classifier was responsible for its high performance, whereas the AdaBoost showed comparatively lower performance because of its sensitivity to noisy samples. Gradient Boosting was able to improve classification ability by sequentially reducing classification errors. Because of the automatic feature extraction process, the CNN model succeeded in obtaining better feature representation and classification accuracy. The fusion of ensemble models with CNN architecture, however, yielded the best overall result, which resulted from the benefits of both models. The rejection mechanism based on the confidence also contributed to the increase in system reliability. In contrast to the predict-for-uncertain-samples approach, the proposed framework discarded low-confidence predictions according to a threshold value. This decreased the chance of mis-classification and enhanced the trustworthiness of decisions.

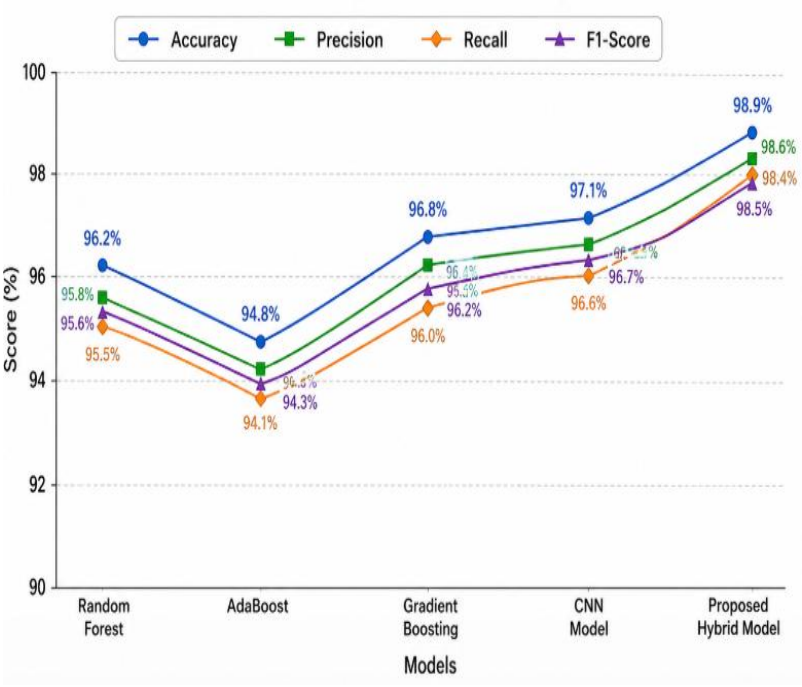
Performance Comparison of Classification Models

Threshold Value	Rejection Rate (%)	Accuracy on Accepted Samples (%)
0.70	2.1	99.0
0.80	4.8	99.4
0.90	8.7	99.8

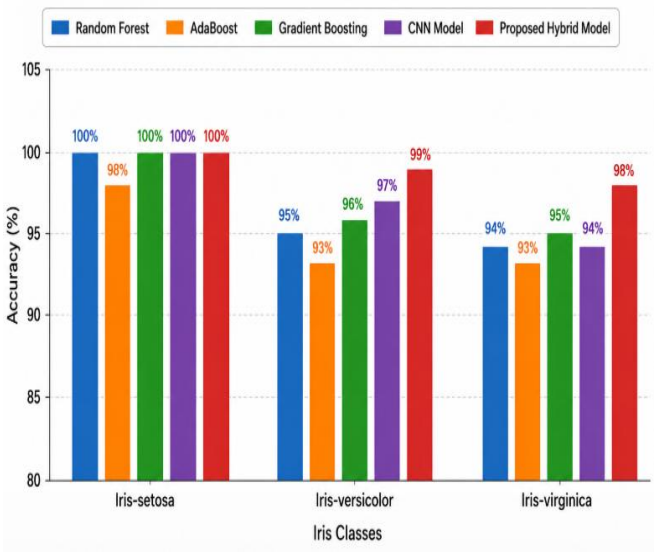
Table 4 shows that the performance of the classification system in terms of reliability of correct classification for accepted samples becomes more reliable with an increase in the rejections threshold. The framework had 99.8% accuracy for accepted predictions and 8.7% rejection rate for the threshold value of 0.90. This demonstrates the trade-off between prediction coverage and reliability. This selective classification is very useful for cyber security applications, where forecasts are not always accurate, and it may be necessary to analyze them manually before making the final decision. The analysis of confusion matrix also demonstrates that the proposed framework achieved less class level misclassification errors in all classes of Iris. The hybrid fusion strategy was able to improve the class separability and narrow down the overlapping regions of the prediction region successfully. CNN component was beneficial in acquiring effective feature learning, and ensemble classifiers had a positive effect on stability and robustness toward variance in classification. Overall, experimental results confirm the effectiveness of the proposed Hybrid Ensemble-CNN framework with rejection mechanism for reliable intelligent classification systems. The framework showed high prediction accuracy, high robustness and high trustworthiness, which makes it a promising solution for future cybersecurity-related applications like intrusion detection, anomaly prediction and intelligent threat analysis.



Accuracy Comparison of Different Models.



Metric Comparison of Different Models.



Class-wise Accuracy Comparison.

The figure 4 presents the classification accuracy results obtained using various Machine Learning and Deep Learning models such as Random Forest, AdaBoost, Gradient Boosting, CNN Model and the proposed Hybrid Model. The proposed hybrid framework with ensemble learning and CNN architecture yielded the best accuracy rate of 98.9%, which shows the benefits of combining two classifiers. The findings show that the hybrid fusion method has better predictive performance than a single method. The line figure 5 displays a comparison of accuracy, precision, recall and F1 score for each classification model. In all the evaluation metrics, the results of the proposed hybrid model always outperformed other models, showing better reliability of the classification model and balanced prediction performance. The graph also shows that ensemble learning based on CNN architecture has a better generalization performance than individual classifiers. This grouped image 6 shows the accuracy of various classifiers by class for the Iris data set which contains three classes: Iris-setosa, Iris-versicolor, Iris-virginica. The proposed hybrid model achieved the best accuracy for all the classes especially for overlapping classes of Iris-versicolor and Iris-virginica. This shows the effectiveness of the hybrid approach in minimizing the misclassification errors at class level.

7. Conclusion and Future Work

In this research, a Hybrid Ensemble-CNN Classification Framework with an intelligent Confidence based rejection mechanism was proposed as a reliable intelligent classification system using Iris dataset as a benchmark dataset. The proposed framework combines ensemble learning models (Random Forest, AdaBoost and Gradient Boosting) with CNN architecture to boost classification accuracy, robustness and predictive reliability. The experimental results showed that the proposed hybrid model achieved higher accuracy, precision, recall and F1 score values than individual classifiers and it successfully rejected uncertain predictions via rejection mechanism. The confidence-based selective classification approach increased the trustworthiness of decisions and can be applied in cybersecurity-related applications like intelligent threat monitoring, anomaly analysis, and intrusion detection. The proposed framework can be further extended with the real-world cybersecurity datasets like NSL-KDD, CICIDS2017, and UNSW-NB15 to validate its efficiency in real life scenarios in future works.

References:

1. C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.
2. T. G. Dietterich, “Ensemble methods in machine learning,” in International Workshop on Multiple Classifier Systems, Springer, 2000, pp. 1–15.
3. L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
4. Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119–139, 1997.
5. A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016.
6. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.
7. G. E. Hinton, L. Deng, D. Yu, G. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, and B. Kingsbury, “Deep neural networks for acoustic modeling in speech recognition,” IEEE Signal Processing Magazine,

vol. 29, no. 6, pp. 82–97, 2012.
8. C. K. Chow, “On optimum recognition error and reject tradeoff,” IEEE Transactions on Information Theory, vol. 16, no. 1, pp. 41–46, 1970.
9. Y. Geifman and R. El-Yaniv, “Selective classification for deep neural networks,” in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4878–4887.
10. Agarwal, G., Katiyar, V., Tiwari, A. (2026). Cloud-Based AI Development Tools: Making Machine Learning is Accessible to Everyone. In: Dutta, D., Das, S., Roy, J., Mahanty, B. (eds) Numerical Optimization in Engineering and Sciences. NOIEAS 2024. Lecture Notes in Networks and Systems, vol 1638. Springer, Singapore.
11. Chaudhary, V., & Tiwari, A. (2025, October). Graph Neural Networks for AWS Cloud Security Analytics: Anomaly Detection Using CloudTrail Logs. In 2025 3rd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT) (Vol. 1, pp. 492-496). IEEE.
12. Tiwari, A., Pandey, R. K., Mishra, S., Singh, A., Singh, S., & Chaudhary, V. K. (2025, August). Machine Learning Pipelines with AWS SageMaker for Customer Churn Analysis. In 2025 International Conference on Sustainability, Innovation & Technology (ICSIT) (pp. 1-6). IEEE.
13. Singh, S., Tiwari, A., Singh, A. K., Singh, H. K., Chaudhary, V. K., & R. Johri. (2025, August) Predicting Life Insurance Risk Using Deep Learning. In 2025 International Conference on Sustainability, Innovation & Technology (ICSIT) (pp. 1-6). IEEE.
14. Tiwari, Ashish , Addimulam, Sushanth Chandra , Dubey, Nilesh , Kumar, K. Kiran , Kapila, Dhiraj & Jain, Kusumlata (2025) Cybersecurity in voice over IP (VoIP) systems : Protecting communication networks from threats, vulnerabilities, and cyber attacks, Journal of Discrete Mathematical Sciences and Cryptography, 28:8, 2989–2999.
15. Tiwari, A., Mishra, S., & Tiwari, S. (2025). AI/ML: Developing Algorithms for Handling Imbalanced Datasets in Classification Tasks. In Computational Intelligence Techniques for 5G Enabled IoT Networks (pp. 17-29). Cham: Springer Nature Switzerland.
16. Srivastava, S., Dubey, N., Thakral, D., Batra, N., Sharma, D., & Tiwari, A. (2024, December). IoT enabled Health Wearables and Monitoring System in Computing. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-6). IEEE.
17. Tiwari, A., Kumar, S., Singh, P. N., Kapila, D., Shukla, A., & Pachori, A. (2024, December). Building an E-commerce Platform with the MERN Stack. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-6). IEEE.
18. Chaudhary, V., & Tiwari, A. (2024, December). Image Dehazing using Fourier Transform and Bilateral Filtering. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-5). IEEE.
19. Mishra, S., & Tiwari, A. (2024, December). Medical Insurance Cost Prediction using AWS Sage-Maker. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-6). IEEE.
20. Singh, N. K., Randhawa, P., Tiwari, A., Lakhanpal, A., & Srivastava, S. (2024, December). Secure and Reversible Data Hiding in ECG Signals for Enhanced Patient Privacy in Telecardiology. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-6). IEEE.
21. Pachauri, M., Rizvi, S. W. A., Rizvi, S., & Tiwari, A. (2024, December). Improving Brain Tumor Detection Using Machine Learning. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-6). IEEE.
22. Tiwari, A., Chauhan, S. S., Singh, S., & Singh, V. (2024, December). Evolution of Cybersecurity in the age of IoT and Cloud Computing. In 2024 International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIA) (pp. 1-4). IEEE.
23. M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, “A detailed analysis of the KDD CUP 99 data set,” in Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009, pp. 1–6.
24. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
25. S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, 4th ed. Pearson, 2020.
26. J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann, 2011.
27. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in International Conference on Learning Representations (ICLR), 2015.
28. F. Chollet, Deep Learning with Python. Manning Publications, 2018.