SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 4559-4575
Open Access. Online & Print

DOI: 10.5281/zenodo.12426504

ENHANCED CYBER THREAT DETECTION IN HEALTHCARE SYSTEMS USING HYBRIDIZED ENSEMBLE MODEL WITH SPIDER WASP OPTIMIZATION

Manujakshi B C[1]*, Shashidhar T M[2], N. Sivakumar[3], Renukadevi S[4]

[1] Associate Professor, Department of Computer Science and Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bengaluru, Karnataka, India.
[2] Professor, Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Harsha Institute of Technology, Bengaluru.
[3] Associate professor, Department of Computer Engineering, Marwadi University, Rajkot, Gujarat,India.
[4] Assistant professor, Department of cloud Technology and Information Security, School of CS&IT, Jain (Deemed-to-be- University), Bengaluru, Karnataka, India.

Received: 08/12/2025
Accepted: 26/02/2026

Corresponding author: Manujakshi B C
(manujakshibc@gmail.com)

ABSTRACT

The modern healthcare industry makes up the digital healthcare industry with data-driven infrastructure for collecting, rather, than patient details from smart sensing devices and providing quick responses for diagnosis, cyber threats become more propensity in making out anomalies from data security and privacy challenges. This is because most traditional cybersecurity models in the healthcare industry are sensitive. A hybridized ensemble ML classifier-based Spider Wasp Optimization (HERFC-SWO) is presented as a threat detection and mitigation enhancement research tool. The model integrates Support Vector Machines (SVM), Random Forest (RF), and XGBoost to improve classification accuracy. It also uses Spider Wasp Optimization (SWO) as a fitness function in the classifier network to maximize feature selection and detection performance. The study will begin with a data collection phase from typical web sources, then engage in Hot-Deck Imputation preprocessing to deal with missing values, and then Z-score normalization to standardize the dataset. Feature extraction will be done using Information Gain (IG) as well as Association Rule Learning (ARL) to select meaningful attributes, with Principal Component Analysis (PCA) applied to reduce dimensionality. To even make the healthcare system more secure, an Advanced Multi-Factor Authentication (AMFA)--based security layer is incorporated into the data transmission across the cloud networks. Validated with various performance measures, the strategy offered will ensure accurate health systems cyber threat detection. Therefore, the system has the highest accuracies (99.13%) in both precision (99.13%) and sensitivity (99.19%), as well as the lowest false positive (2.49%) and false negative (3.54%).

KEYWORDS: Hybridized Ensemble ML Classifier, Spider Wasp Optimization, Healthcare Cybersecurity, Threat Detection, Feature Selection, Multi-Factor Authentication

1. INTRODUCTION

The ever-changing landscape of cyber threats, as seen by the constant advent of new attack methods, poses a growing threat to the integrity of sensitive and confidential information in the global digital realm [1]. As a result, cybersecurity has emerged as a critical concern, necessitating the implementation of robust defense mechanisms for ICT systems and applications. Technological and operational advancements have caused a substantial change in the field of cybersecurity as applied in the world of business [2]. Supported by evidence from various studies, cybersecurity thus becomes a fundamental factor in determining business success or failure. Breaches in cybersecurity greatly affect the market valuation, reputation, and competitive positioning of an organization [3, 4]. These breaches place quite a heavy financial burden on organizations and individuals alike. Information security spending has substantially increased over the last few years. The recent fast digitization of the healthcare sector has resulted in significant improvements in patient care and operational efficiency [5]. However, this increased reliance on digital technology has exposed the business to significant cybersecurity concerns, including data breaches, ransomware attacks, and unauthorized utilization of sensitive medical information [6]. Cyber threats to healthcare will interrupt critical services, compromise privacy, and result in financial losses. To deal with such threats, predictive analytics based on Artificial Intelligence (AI) and Machine Learning (ML) is emerging as a proactive way of identifying and mitigating cyber threats before they incur substantial damages [7].

Organizations that are working in the healthcare sectors and other healthcare organizations are finding the hostile environment for data security fast becoming rather dangerous with the increase in the number of cyber-security incidents. These situations have forced the healthcare sector to allocate large resources toward protecting their systems [8]. A major breach brings other threats to bear, such as those on medical equipment and infrastructure, data corruption, and financial fraud [9]. To this end, many governance measures have been instituted to ensure that the highest standards are kept to safeguard hospital electronic infrastructure against the consequences of cybersecurity attacks [10]. The majority of research work has disclosed the correlation between information system cybersecurity and the performance of organizations. Thus, a viable framework for predictive analytics in cybersecurity in the healthcare system must integrate many layers of security [11, 12]. These include intrusion detection systems (IDS), behavioral analysis, encryption techniques, and blockchain-based access control mechanisms. AI-based security solutions evolve rapidly to accommodate changes in threats for the betterment of realizing effective use of resources in cybersecurity management [13]. Besides, reinforcement learning and hybrid optimization gain significance as possible avenues to improve threat detection accuracy by reducing false positive incidences during real-time decision-making [14]. Predictive models help in proactively detecting malicious activities along with automated response mechanisms targeting the betterment of security protocols [15]. This, therefore, reduces existing vulnerabilities of the system hence providing timely and constant protection against agile and ever-changing cyber threats.

Thus, this is going to define an AI-ML predictive analytics framework for cyber security within the healthcare industry. Specifically, this researcher proposes the HERFC-SWO method for improved identification of cyber threats within healthcare systems. It is a combination of RF, SVM, and XGBoost to maximize classification accuracy, bearing in mind the increasing incidence of ransomware attacks. The SWO method is also used as a fitness function to enhance attack detection performance and optimum feature selection. In the advanced security layer, an AMFA protocol is included to enhance highly secure data transfers over cloud networks. The framework aims at improving incident response efforts, automating risk assessment, and improving threat predictability. Such work will help improve healthcare security resilience because it will utilize AI and ML algorithms to warrant the availability, confidentiality, and integrity of crucial patient data. It directly reduces any future possibility of a cyber-threat while promoting security and trust in digital healthcare services. Some of the contributions his model brings into existence are

- Develop a robust cyber threat detection model for healthcare systems by integrating ensemble ML classifiers with SWO.
- Enhance feature selection and classification accuracy using IG, ARL, and PCA.
- Improve preprocessing efficiency by utilizing Hot-Deck Imputation for missing data handling and Z-score normalization for standardization.
- Strengthen security mechanisms in healthcare data transmission through the implementation of an AMFA protocol.
- Validate the proposed HERFC-SWO strategy over experimental analysis, comparing its performance with existing cybersecurity models.

2. RELATED WORKS

By examining how cybersecurity affects healthcare workers' use of electronic health record systems and how it affects their performance, Ramayah et al [16]

create a novel strategy. By combining cybersecurity, psychological trust, and a technology acceptance model (TAM), it goes beyond conventional adoption theories and concentrates on the real-world use of EHR systems and how they affect performance. A thorough understanding of the dynamics at play is offered by the analysis of data from 459 medical professionals using Structural Equation Modeling-Artificial Neural Networks (SEM-ANN). The results imply that the influence of cybersecurity on system efficiency may differ depending on the situation.

Threats are analyzed and predicted by Abel et al [17] to enhance cyber supply chain security. Additionally, ML methods were used in conjunction with cyber threat intelligence (CTI) to analyze and forecast attacks based on CTI features. This makes it possible to determine the intrinsic CSC vulnerabilities so that the right control measures can be implemented for the enhancement of cybersecurity as a whole. The Microsoft Malware Prediction dataset is used to create predictive analytics utilizing several ML methods and CTI data to illustrate the applicability of our methodology.

To improve threat intelligence, Nagamalla et al [18] present an extensive structure for industrial Internet of Things (IoT) cybersecurity that integrates several methods. Raw IoT data is refined by the data preparation and integration program using a painstaking 20-step procedure. While the RF algorithm concentrates on statistical modeling for proactive threat identification, the time series analysis algorithm explores temporal trends. By identifying abnormalities and capturing temporal relationships, the LSTM Ensemble method expands the study into behavioral ML. Predictive analytics and behavioral model results are combined via the Weighted Average Ensemble, which takes use of their correlation to improve threat intelligence.

To secure, prepare, and modify the healthcare system to handle future pandemics, Petar et al [19] expand on the understanding of how to instruct and train new AI systems. The main goal is to create a prototype healthcare system that uses edge health devices providing real-time data and is backed by autonomous artificial intelligence. To (1) self-optimize predictive cyber risk statistics of healthcare provider failures and (2) self-adaptive predicting of medical manufacture and supply chain bottlenecks during future pandemics, the article builds two case circumstances for integrating cybersecurity with autonomous artificial intelligence. Using optimization techniques, the testing scenarios are designed to address the logistical difficulties and interruption of intricate production and supply chains for vaccine delivery.

Jyri et al. [20] offer a cutting-edge master's degree with three main focus domains: digital skills, societal skills, and health sector skills. Because cybersecurity and healthcare innovation are intertwined, the ManagiDiTH curriculum incorporates cybersecurity education into a larger framework of digital transformation capabilities for health professionals. The ManagiDiTH architecture tackles the unique cybersecurity issues that the healthcare industry faces, such as ransomware attacks, data breaches, and the growing use of technology in patient care. The study introduces a new educational program that integrates cybersecurity digitalization elements of important healthcare fields.

Archana et al. [21] map different visual representations using a key learning network built on the ResNet-50 architecture. To decrypt the encrypted image, we initially transform it back into its "plaintext" form using reconstructive systems. By leveraging the user's local information setting, data mining may be easier, and a Return on Investment (ROI) architecture can be offered whenever the hidden entities are finally discovered. The results of the security evaluation and the thorough empirical setup imply that the proposed approach might produce an output with an unprecedented degree of security and power.

To speed up the retrieval procedure and increase patient data security, Arunprasath et al. [22] suggest a simplified ontology-based retrieval technique for healthcare multimedia data. An approach for extracting pertinent data from huge datasets was an ontology-based retrieval system. In hierarchical learning, learning tasks were arranged in a hierarchy where lower-level policies were guided by higher-level policies. Multiple models were merged to enhance a reinforcement learning technique's performance through ensemble model-based reinforcement learning. With a 92% accuracy rate, the Simplified Ontology-Based Retrieval Algorithm demonstrated exceptional performance and was very successful in information retrieval tasks.

3. PROBLEM DEFINITION

Technological development in Information and Communication Technology (ICT) has facilitated accessibility, communication, and operational efficiency in the last 50 years, but these technologies have also placed enterprises under serious cybersecurity threats. With the increment in adverse activities in the cyberspace arena, the security of data has become a pressing issue in cases wherein sensitive information regarding patient health is concerned, such as in health care [23]. Hospitals today find themselves facing multiple cybersecurity threats, such as data breaches, ransomware, phishing, and IoT vulnerability. Such threats threaten the integrity of medical infrastructure, data loss, and financial fraud. In defending against these threats,

healthcare institutions are applying security protocols and working toward meeting the unique regulatory mandates provided in policies like HIPAA and GDPR. Outside research indicates that a correlation exists between cybersecurity preparedness and improved organizational performance, suggesting that a strong security posture enhances a firm's reputation and profitability while increasing the resiliency of its overall systems [24]. It is compliance-driven cybersecurity management that reduces risk and increases operational security across the board.

4. PROPOSED METHODOLOGY

Initially, healthcare data is collected from standard web sources, and then the online data analysis layer is activated to proceed further with the data preparation process. In data preprocessing, Hot-Deck Imputation dealt with missing values by selecting donor cases that best matched the recipient scenario, preserving empirical deviations and associations better than mean imputation. Besides, Z-score normalization ensures the dataset has a mean of zero and a standard deviation of one so that statistical analysis can be improved. IG and ARL would suffice for an appropriate feature extraction process to identify significant attributes that contribute to the classification. Furthermore, PCA is used to reduce dimensionality by converting correlated data into uncorrelated feature spaces. With the tremendous increase in the complexity and frequency of cyber threats, particularly those involving ransomware, it is critical to establish powerful and responsive security procedures. As a result, our study suggests that HERFC-SWO can be used to identify and mitigate intrusions more effectively. The suggested model improves prediction accuracy by combining RF, SVM, and XGBoost classifiers. Furthermore, the SWO algorithm is used as a fitness function in classifier networks to improve feature selection and threat detection performance. In addition, the AMFA protocol is implemented in the security layer to reinforce data security in cloud networks. This protocol thus improves authentication mechanisms with multiple verification techniques for enhanced resilience against unauthorized access and data breaches. The architecture of the developed model appears in Fig. 1.

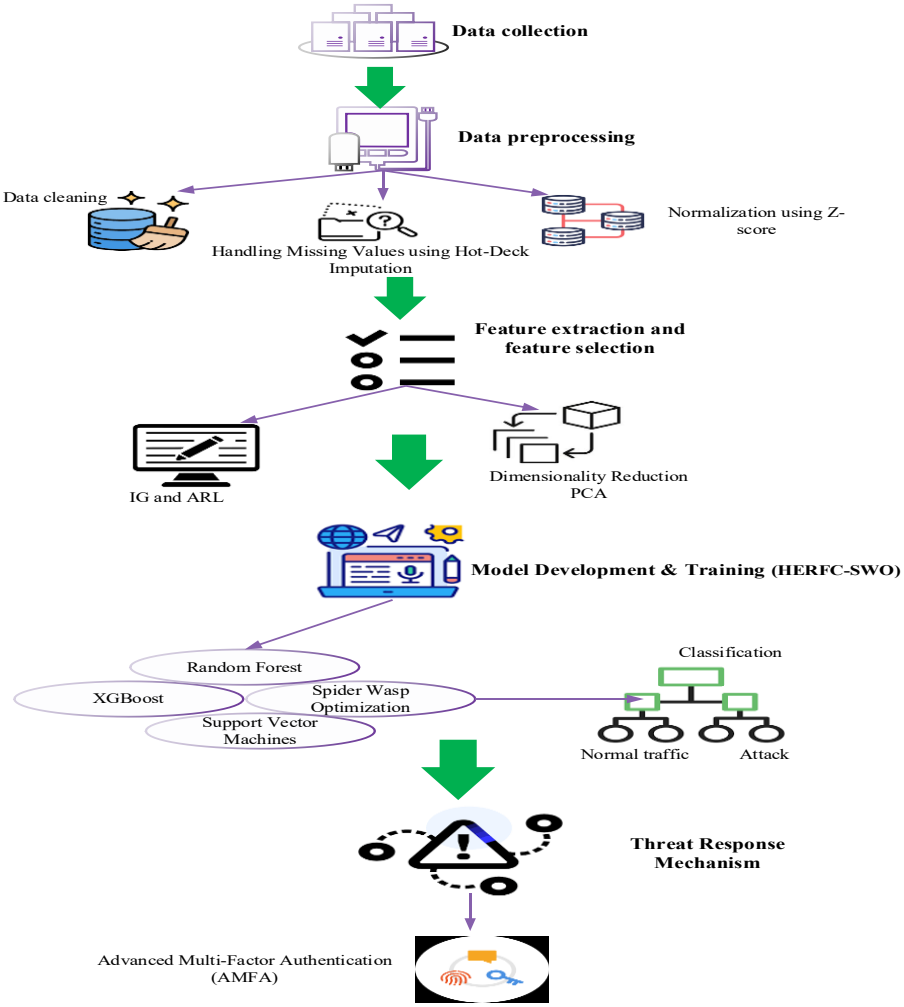

Figure 1: Developed technique architecture

4.1. Data collection

The CSE-CIC-IDS2018 dataset [25], established by the University of New Brunswick, comprises a complete intrusion detection dataset for DDoS (Distributed Denial of Service) attacks. This dataset has multiple network traffic logs captured in 2018; thus, this dataset is often applied for cybersecurity research, especially about IDS based on ML. The dataset includes multiple corresponding CSV files for various days, forcing researchers to preprocess and even balance before effective analysis. The dataset has features totaling up to 80, including traffic ranges, such as destination port, protocol type, flow duration, and total forward and backward packets received, among other statistical network flow attributes, the most important of which is "Label" that determines whether the current network flow is malicious or benign. The dataset's structure replicates real-world network behavior, including both regular and DDoS assault traffic, making it an important resource for developing effective intrusion detection models.

4.2 Preprocessing

Data cleaning: This includes controlling the dataset's noise, outliers, and missing values. Cleaning ensures data quality, as malware databases contain inconsistent or missing data.

- **Handing missing values - Hot-Deck Imputation**

The word "decks" of computer cards alludes to the hot-deck imputation method [26]. When employed in the context of missing data imputation, these decks contain observations from various cases, known as "donor cases," that match the "recipient" scenario on a set of specified variables from the same data collection. The inaccurate value is imputed by using either an exact match or a randomly picked measured value from the hot deck. The procedure is repeated numerous times to provide double imputed data, which has the extra benefit of representing uncertainty for both sampling and missing data if the donor is picked at random from the collection of matches. The acceptable values from the observed distribution are used as donors in hot-deck imputation, which has an advantage over mean imputation in that empirical deviations and relationships are better preserved.

In line with the Hot Deck concept, Each absent value x^* of a variable X (x^* as a component of X_{mis}) is ascribed using a recorded value x (x as a component of X_{obs}). Now, the individual with the absence is frequently referred to as the "recipient," The person who provides the imputation value is referred to as the "donor". In general, one should distinguish between two methods: Random Hot Deck operations randomly select a value for the donor from an underlying donor pool.

- **Z-score normalization**

It outlines the technique for normalizing each value in a dataset so that the standard deviation is one and the mean is zero in equation (1).

$$Z = \frac{(X-\vec{\mu})}{\vec{\sigma}} \quad (1)$$

X is shown as the initial value, $\vec{\mu}$ is represented as the data mean, and is regarded as the data's standard deviation. During the feature extraction stage, the normalized dataset is improved to yield statistical and dynamic properties.

4.3 Feature extraction

To assess the valuable information recovered from the cleaned dataset, two methodologies are used: IG and ARL. When it comes to feature extraction, IG and ARL function effectively together to identify relevant qualities that contribute significantly to the association rules.

- **Information Gain (IG)**

The importance of a feature for forecasting the target variable is assessed using IG [27]. When a specific attribute is known, it calculates the decrease in entropy or unpredictability of the target variable. IG is an entropy-based feature evaluation technique that quantifies how much information a feature contains about the target class. IG can find the features that contain the most information based on the target class. Features with a high IG are usually chosen since they are strongly related to the target class and produce the best classification results. Nonetheless, Instagram is unable to remove unneeded features.

The amount of information available both before and after the attribute value was known typically impacts how much knowledge is acquired. For numerous classes, IG has a maximum value of one. Equation (2) specifies the formula for entropy analysis of more than two classes.

$$G(Y) = \sum_{i=1}^{k} Q(y_i) \log_2 Q(y_i) \quad (2)$$

Let, Q is denoted as the number of classes. Moreover, feature Y of I_G and the class labels Z is designed in eqn. (3) and (4).

$$I_G(Y,Z) = G(Y) - G(Y/Z) \quad (3)$$

$$G(Y/Z) = -\sum_j Q(z_i) \sum_i Q(y_i/z_i) \log_2(Q(y_i/z_i)) \quad (4)$$

Let, (Y/Z) is denoted as the entropy of Y and $G(Y/Z)$ is considered as entropy of Y after seeing Z. Since I_G is considered a filter technique, when dealing with big multidimensional data, it scales effectively.

- **PCA**

PCA is a statistical procedure that uses an orthogonal change. A set of correlated variables is converted into a set of uncorrelated variables using PCA [28]. For exploratory data analysis, PCA is used. Furthermore, PCA can be used to examine the relationships between a group of variables. It can therefore be used to reduce dimensionality.

Let's say a dataset $z^{(1)}, z^{(2)}, z^{(3)}, \ldots . z^{(n)}$ has η n-dimension data that must be transformed into dimension inputs i -dimension $(i << n)$ using PCA. Below is a description of PCA:

Data Standardization: Equation (5) states that the raw data must have a unit variance & a zero mean.

$$k_j^i = \frac{k_j^i - \overline{k_J}}{\sigma_j} \forall_j \quad (5)$$

Using equation (6), determine the co-variance vector of the raw data.

$$\sum = \frac{1}{n} \sum_i^n (k_j)(k_i)^t, \sum \in E^{n*n} \quad (6)$$

Utilizing the formula from equation (7), determine the covariance matrix eigenvector and eigenvalue.

$$v^t \sum = \gamma \mu$$

$$V = \begin{bmatrix} . & . & . \\ v_1 & v_{2\ldots.} & v_n \\ . & . & . \end{bmatrix}, v_i \in E^n \quad (7)$$

It is necessary to project raw data onto a n -dimensional subspace: Top n The covariance matrix's eigenvector is selected. Equation (8) provides the necessary vector calculation.

$$Z_i^{new} = \begin{bmatrix} v_1^t z^i \\ v_2^t z^i \\ \ldots \ldots . \\ \ldots . \ldots \\ v_k^t z^i \end{bmatrix} \in E^k \quad (8)$$

Thus, if the unprocessed data is with η dimensionality, it'll be condensed into new k data presented in three dimensions.

4.4 Model Development and Training

The proposed method is the novel HERFC-SWO strategy. The study encompasses three highly efficient ML models, XGBoost, RF, and SVM, where each of the classification models offers its own merits mostly for attack detection. Such models are widely preferred owing to robustness, efficiency, and the capability to address highly complicated datasets. An elaborate discussion of each of the models is presented below about their working principles and mathematical fundamentals.

- **XG Boost**

XGBoost (Extreme Gradient Boosting) [29] is an ensemble tree method that straightforwardly optimizes weak learners using a gradient descent framework. It works on the principle of boosting where a weak learner gradually improves itself by iteratively correcting its older version in the ensemble. Each time new models are trained to give higher weightage to those instances that could not be classified in the previous iterations of the boosting, resulting in a highly optimized final model. The data is first fitted onto the weak classifier by XGBoost. Subsequently, the data is fitted to an additional weak classifier to improve i^{th} Accuracy without modifying the existing model. The procedure is repeated until the highest accuracy is achieved. When selecting the conditional function in XGBoost, the objective function takes the loss function and normalization term into account. Equation (9) provides the XGBoost goal function.

$$O_f = \sum_{i=1}^m K(\hat{x}_i, x_i) + \sum_{i=1}^n E(f_i) \quad (9)$$

Let, K is demoted as a loss function measuring the difference between predicted and actual values, $\hat{x}_i$ is considered as a predicted label, x_i is denoted as an actual label, $E(f)$ is denoted as the regularization term that penalizes the training tree's computational function. By minimizing this objective function, XGBoost efficiently finds the optimal decision tree structure while controlling model complexity.

- **Random Forest (RF)**

RF [30] is an ensemble learning technique that combines many Decision Trees to increase classification accuracy and reduce overfitting. It builds a lot of decision trees during the training phase and takes the average prediction of all trees to make the final call. In RF the output is the mean prediction of all the tree classifiers. Several decision trees are constructed and combined using a RF ensemble classifier to give the best output. A large number of decision trees are created using bootstrapped samples of the dataset. Each decision tree is trained

independently on a random subset of features. An average (for regression) or majority voting (for classification) is taken as the final prediction across all the trees.

Let the given information X with a minimum of $v=1$ and a maximum amount of V : The sample's forecast x' is calculated by averaging the forecasts. $\sum_{v=1}^{V} g_v(x')$ from each tree for that is demonstrated in eqn. (10).

$$k = \frac{1}{V}\sum_{v=1}^{v} g_v(x') \quad (10)$$

Let, $g_v(x')$ be the prediction of an individual decision tree, and V be the total number of trees in the forest.

- **Support vector machine (SVM)**

SVM [31] is a supervised learning model that is popular for classification and regression analysis. It tries to find the hyperplane that best differentiates between the data points of different classes. For the linearly separable data, SVM creates the decision boundary with maximum margin, while for the non-linearly separable data, it applies kernel tricks to transform the input space. Given a labeled training dataset, SVM finds a hyperplane that maximizes the margin between two classes. If the data cannot be separated linearly, SVM uses kernel functions (such as polynomial or RBF) to map the data into a space of higher dimension where a hyperplane can be drawn. The optimal hyperplane is determined based on maximizing the margin, which represents the distance between the hyperplane and the nearest data points (support vectors).

For a binary classification problem with training samples (Z_i, X_i), where Z_i is the feature vector and X_i is the class label $X_i \in \{-1,1\}$ The optimal hyperplane is defined as Eq. (11)

$$W_t Z + B = 0 \quad (11)$$

Where: W is the weight vector that defines the hyperplane's orientation. Z is the input feature vector. B is the bias term.

The classification decision function is given by Eq. (12)

$$F(Z) = sgn(W_t Z + B) \quad (12)$$

To maximize the margin, SVM minimizes the objective function by Eq. (13)

$$\frac{1}{2}\|W\|^2 \quad (13)$$

subject to Eq. (14)

$$X_i(W_t Z_i + B) \geq 1, \forall_i \quad (14)$$

If data is not linearly separable, a slack variable ζ is introduced, and the optimization problem becomes in Eq. (15)

$$min\frac{1}{2}\|W\|^2 + d\sum_{i=1}^{n}\zeta_i \quad (15)$$

where d is the regularization parameter controlling the trade-off between maximizing the margin and minimizing misclassification.

The HERFC-SWO strategy is devised for attack classification, fusing three ML models, namely XGBoost, RF, and SVM. Optimization for the models is achieved using SWO. XGBoost is a boosting algorithm that iteratively improves weak classifiers for enhanced classification. RF adds robustness by training many decision trees that are combined, producing an average of their predictions to counteract overfitting. SVM identifies optimal decision boundaries for attack detection using hyperplane separation. SWO optimization fine-tunes hyperparameters for improved detection accuracy and efficiency. This ensemble method guarantees high performance for distinguishing abnormal from attack traffic in cybersecurity systems. Then, the fitness function of SWO was fed directly into the classifier networks for attack detection. The architecture of the HERFC-SWO model is shown in Fig. 2.

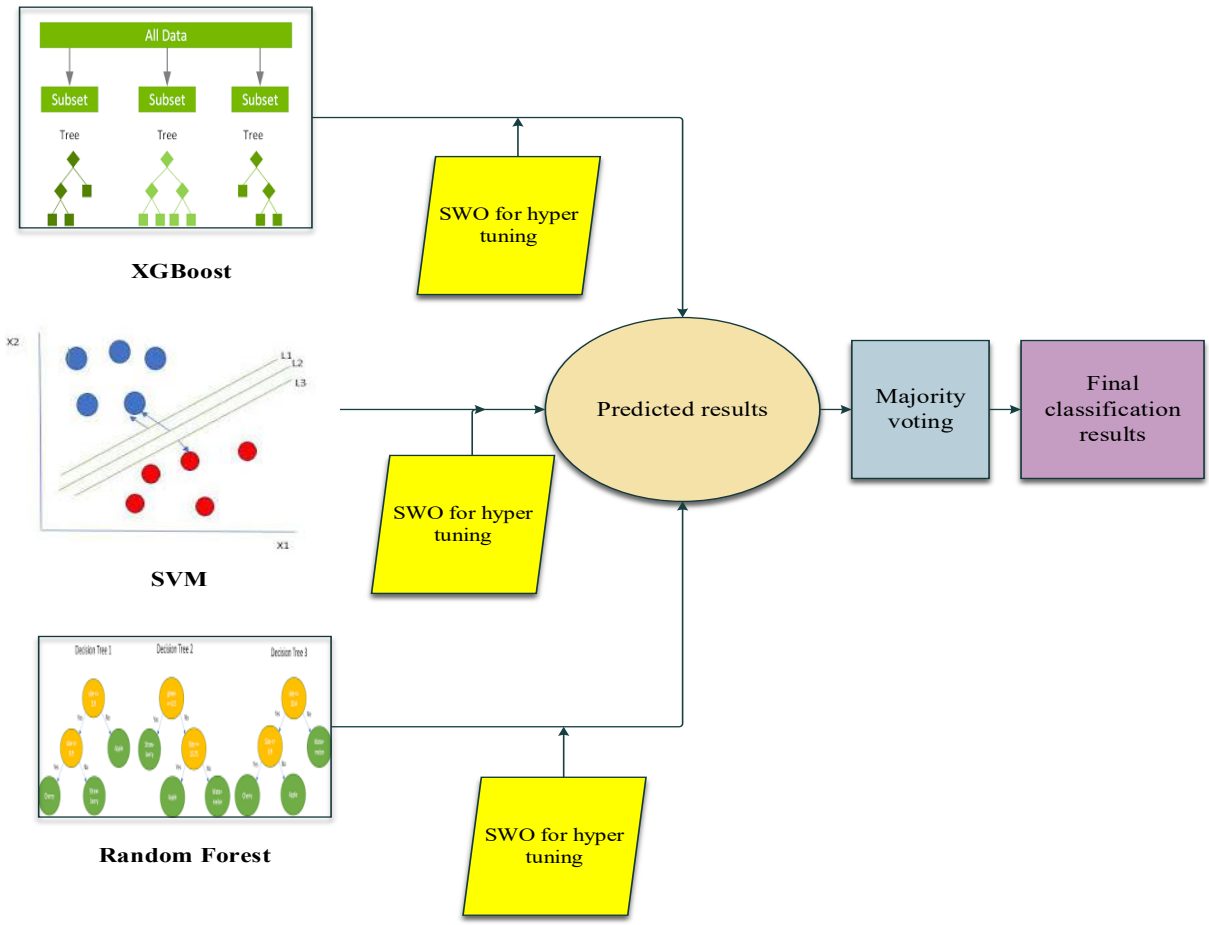

Figure 2: Design of HERFC-SWO model

4.4.1. Spider wasp optimizer (SWO)

The SWO [32] is a bio-inspired metaheuristic algorithm applied for hyperparameter tuning of ML models. It starts by defining the search space where hyperparameters are selected for models such as SVM, RF, and XGBoost. A population of wasps representing candidate hyperparameter sets is randomly initialized inside the search space. The optimization process consists of several phases. In the Searching Phase, wasps explore the search space randomly, examining various hyperparameter configurations in search of better-performing solutions. Then, in the Following and Escaping Phase, promising solutions gather other wasps, which narrow the search down to the regions that shine best. During the Hunting and Nesting Phase, the best ones are kept, while weaker ones are replaced to avoid stagnation through a Lévy Flight mechanism. The Mating Phase introduces crossover operators to form new hyperparameter combinations and hence maintain diversity in the search. Each candidate set is evaluated against a fitness function that balances accuracy and computational efficiency. The optimization process continues until it meets the convergence criterion, which may be a maximum number of iterations or minimal improvement in the fitness value. Eventually, the best hyperparameters found by SWO will be adopted to train the ML model for attack detection. Below is a detailed stepwise explanation of how SWO can be applied to optimize ML models.

Define the Hyperparameter Search Space: Before using SWO, define the hyperparameters of the ML model that must be tuned. For each hyperparameter, the search space has either continuous or discrete variables.

- **SVM:** Kernel type, C (regularization parameter), gamma.
- **RF:** Number of trees, max depth, min samples split.
- **XGBoost:** Learning rate, max depth, number of estimators

Each wasp in the population represents a vector of hyperparameters in the search space.

Initialize the SWO Population: SWO starts with an initial population of female spider wasps, where each wasp represents a candidate hyperparameter set. The position of each wasp (solution) is initialized randomly within the defined bounds. The population is spread across the search space to ensure a diverse set of initial solutions. The best-performing wasps (solutions) are selected based on their fitness

function.

Mathematically, the initial population can be represented as Eq. (16)

$$s_w(i) = l + R \times (h - l) \quad (16)$$

where: $s_w(i)$ is the initial position of the i^{th} spider wasp, l and h are the lower and upper bounds of the search space, and R is a random number in the range [0,1]

Searching Phase (Exploration of Search Space): This stage mimics the female wasp behavior to look for potential spiders to feed their offspring. Each wasp moves randomly through the search space to find a better hyperparameter combination. Movements are influenced by a constant step size and random normal distribution. The candidate wasp evaluates different locations and chooses the one that maximizes the fitness function. Mathematically, the new position of a wasp is given by Eq. (17).

$$S_w^{t+1}(i) = S_w^t(i) + \delta_1 \times (S_w^t(x) - S_w^t(y)) \quad (17)$$

where: $s_w{}^t(i)$ is the position of the i^{th} wasp at iteration t, $s_w{}^t(x)$ and $s_w{}^t(y)$ are two randomly selected wasps, δ_1 is a motion step computed as Eq. (18).

$$\delta_1 = R_n + R_1 \quad (18)$$

where Rn follows a normal distribution and $R1$ is a random number in the range [0,1]. Ensures comprehensive exploration of the hyperparameter space and lowers the likelihood of becoming stuck in local optima.

Following and Escaping Phase (Exploitation of Best Solutions): The spider runs away to escape the wasp, but instead, it gets captured by the wasp following it. If a wasp finds a good hyperparameter set (better solution), it follows that solution. The best solutions attract other wasps enhancing their search in the high-performing region. The exploitation process decreases randomness gradually to fine-tune the best solution. Mathematically, the update rule is:

$$S_w^{t+1}(i) = S_w^t(i) + D \times \left(s_w(bst) - S_w^t(i) + R_5 \times nor(0, K)\right) \quad (19)$$

where: $s_w\ (bst)$ is the best hyperparameter set found so far, D is a distance-controlling factor, initialized at two and gradually decreasing to 0, R_5 is a random vector, adding stochasticity, K is a dynamic range parameter, reducing step size over iterations. It directs the population toward the best solutions, eliminates needless exploration, and assures convergence to the ideal hyperparameter values.

Hunting and Nesting Phase (Refinement): This is the phase in which a wasp nests, building a nest and storing its prey in it. The "nests" hold the best hyperparameter sets. New, better solutions displace the worst. The mechanism of Levy Flight introduces small disturbances to halt stagnation.

Mathematical relation is detailed in Eq. (20).

$$S_w^{t+1}(i) = s_w(bst) + \alpha \times 1v \quad (20)$$

Let lv be the levy Flight introduces small, random jumps to explore nearby solutions, and α be the scaling factor that controls the magnitude of jumps. It prevents early convergence. Also enhances population diversity.

Mating Phase (Crossover for New Solutions): This phase introduces crossover operations to create new hyperparameter combinations. The best-performing hyperparameter sets exchange information using crossover. The crossover rate (co) controls how much information is exchanged. The mathematical relation is detailed in Eq. (21). From their general location functions, the best solutions are taken forward for future iterations. The weakest are replaced. This is continued until a stopping criterion is reached, which can be reaching the maximum number of iterations, or where further improvement in the fitness values dips below a specific threshold co.

$$s_w(new) = cr \times s_w(f) + (1 - cr) \times s_w(m) \quad (21)$$

where: $s_w\ (f)$ and $s_w\ (m)$ are female and male wasps (parent solutions), cr is the crossover probability. It introduces new hyperparameter sets and enhances exploration by combining strong solutions.

Evaluate Fitness (Hyperparameter Performance): Each set of hyperparameters is evaluated using a fitness function based on Eq. (22).

$$f(s_w) = val\_acc - \beta \times comp\_cost \quad (22)$$

where β balances accuracy and resource efficiency. The best hyperparameters are those that achieve high accuracy with minimal overhead.

Updating the Population and Convergence: From their general location functions, the best solutions are taken forward for future iterations. The weakest are replaced. This is continued until a stopping criterion is reached, which can be reaching the maximum number of iterations, or where further improvement in the fitness values dips below a specific threshold.

Selecting the Best Hyperparameters: Thus, the

optimum hyperparameters identified by SWO are selected for the final training of the ML model after the optimization process. The Pseudocode of SWO to hyper-tune HERFC parameters is detailed in algorithm:1

```
Start
{
Initialize the population of wasps with random hyperparameter sets
Evaluate the fitness of each wasp using a defined fitness function
While stopping criteria not met (e.g., max iterations or convergence threshold):
Searching Phase (Exploration)
  For each wasp:
  {
    Compute new position using motion step and normal distribution
    Evaluate fitness of new position
    If the new position is better, update Wasp's position
  }
Following & Escaping Phase (Exploitation)
  For each wasp:
  {
    Identify best-performing wasps
    Adjust position toward best solutions using distance-controlling factor
    Reduce randomness to fine-tune search
  }
Hunting & Nesting Phase (Refinement)
  For each wasp:
  {
    If a wasp is among the worst solutions:
      Apply Lévy Flight to introduce perturbations
      Replace weak solutions with better ones
  }
Mating Phase (Crossover)
  For each wasp:
  {
    Select parent wasps based on fitness.
    Apply crossover with probability.
    Generate offspring with mixed hyperparameters.
  }
Update Population & Check Convergence
Evaluate the fitness of new wasps
Retain best solutions and discard weak ones
If improvement in fitness is below the threshold, stop iteration
Select the best hyperparameters found by SWO
Train ML model using optimized hyperparameters
}
End
```

4.5. Enhance security using AMFA protocol

In addition, an AMFA protocol is integrated into the security layer as additional protection for the site against AMFA during transcription over the cloud networks to the data. Increased dependence on the cloud for healthcare systems has necessitated powerful security mechanisms to protect sensitive patient information from cyber threats and unauthorized access, which includes the integration of the AMFA protocol with the security framework. AMFA improves security because it includes multiple factors for authentication based on knowledge (e.g., password, PIN), possession (e.g., OTP, smart card), biometrics (e.g., fingerprint, face), and location- or device-approach (e.g., GPS location, device reputation). Mathematically, the authentication probability $p(a)$ can be represented as the joint probability of successful verification across all factors using Eq. (23).

$$p(a) = p(kba) \times p(pba) \times p(ba) \times p(caa) \quad (23)$$

where each term represents the probability of successful verification for a specific authentication factor. Furthermore, the security is made much better by including risk-based authentication (RBA), which ensures the agility feature of adapting security requirements as per users' behavioral pattern

analysis. The risk score r is obtained through a weighted sum of the anomaly-detection factors, using the following formula in Eq. (24).

$$r = W_1F_1 + W_2F_2 + W_3F_3 + \ldots \ldots W_nF_n \quad (24)$$

where F_i represents different risk factors such as failed login attempts, unusual login location, and device fingerprinting, with corresponding weights W_i. If r exceeds a predefined threshold t, an additional authentication factor is triggered. For secure transmission on the cloud, AMFA is interfaced with end-to-end encryption (E2EE) and blockchain-based identity management in preventing unauthorized access and tampering. AMFA implementation will strengthen privacy, integrity, and confidentiality in digital healthcare infrastructures. It will therefore ensure resilient protection from all cyber threats.

5. RESULTS AND DISCUSSION

The proposed technique used the CSE-CIC-IDS2018 dataset and is evaluated against prevailing ML models based on several key performance metrics, including accuracy, sensitivity, precision, F1-score, false positive rate (FPR), false negative rate (FNR), and specificity. The designed technique is implemented in the python tool, and the simulation setup is detailed in Table 1.

Table 1: Simulation setup of the developed model

Parameter	Value/Description
Programming Language	Python
Simulation Tool	Google Collab
Hardware Configuration	Intel i7/i9, 16GB RAM,
Dataset Used	CSE-CIC-IDS2018 dataset
Data Preprocessing Methods	Hot-Deck Imputation, Z-score Normalization
Feature Selection Methods	IG, PCA
ML Models	XGBoost, RF, SVM
Hyperparameter Optimization	SWO
Optimization Algorithm Parameters	Population size = 50, Max iterations = 100
Performance Metrics	Accuracy, Precision, Recall, F1-score, AUC-ROC
Evaluation Method	10-fold Cross-Validation
Training-Testing Split	80% Training, 20% Testing

5.1 Comparison Analysis

The comparisons between different optimization techniques applied to IDS, where different ML models were evaluated on the basis of accuracy, precision, sensitivity, F-measure, specificity, FPR, and FNR. The results show the proposed model's superiority over existing techniques along all major evaluating metrics is shown in Table.2.

Table 2: Performance comparison of existing techniques

Optimization techniques	Accuracy	Precision	sensitivity	F-measure	Specificity	FPR	FNR
RF	0.85349	0.85345	0.85396	0.85406	0.85398	0.04267	0.049528
SVM	0.88746	0.88741	0.88783	0.88786	0.88781	0.03946	0.04612
XGBoost	0.91924	0.91918	0.91971	0.91974	0.91968	0.03512	0.04367
LR	0.93512	0.93507	0.93562	0.93569	0.93559	0.03361	0.04259
KNN	0.94853	0.94848	0.94887	0.94896	0.94884	0.03064	0.03986
SEM-ANN	0.95761	0.95757	0.95781	0.95794	0.95779	0.02951	0.03837
Proposed	0.99138	0.99130	0.99192	0.99200	0.99189	0.02496	0.03548

The graphical representation of the developed technique performance with existing techniques is detailed in figs. (3-9),

Accuracy specifies the ratio of instances correctly classified. Of the traditional models, among KNN (94.85%) and SEM-ANN (95.76%), few were able to surpass RF (85.34), SVM (88.74), and XGBoost (91.92). Certainly, the proposed technique would obtain better classification capability and higher robustness in normal and malicious traffic distinction.

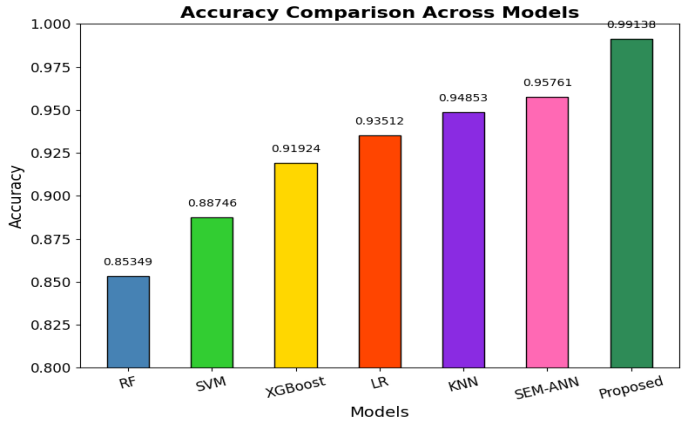

Figure 3: Accuracy analysis

Precision is the ratio of correctly predicted attack traffic out of total attacks. Higher precision indicates fewer false alarms. While RF (85.34%) and SVM (88.74%) had moderate precision, XGBoost (91.91%) and LR (93.51%) both showed better measures than RF and SVM. KNN (94.85%) and SEM-ANN (95.75%) further improved precision, showing reduced misclassification of attack traffic. The highest precision (99.13%) was achieved by the proposed model, which indicates the superior ability to minimize false alarms and accurately detect attacks.

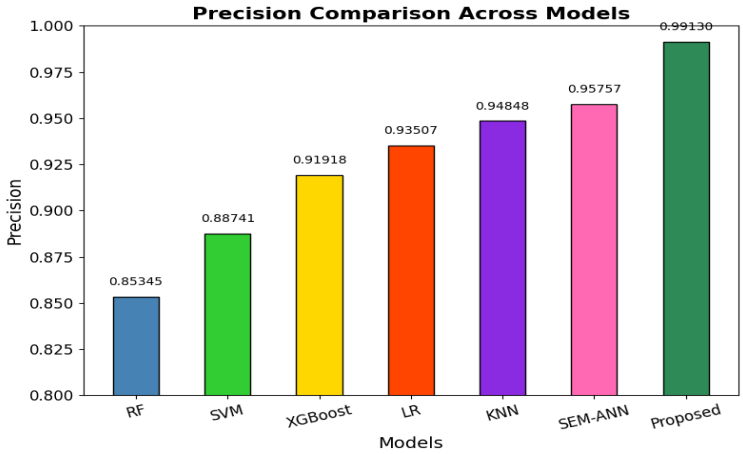

Figure 4: Precision analysis

Sensitivity (or recall) defines an actual attack traffic instance identified by the model as attack traffic. RFs (85.39%), SVMs (88.78%), and XGBoost (91.97%) have shown moderate recall. LRs (93.56%) and KNN (94.88%) had higher sensitivity. The recall was further improved by 95.78% of SEM-ANN, which would capture most attack instances. The proposed model has the highest recall (99.19%), ensuring that almost all instances of attack are detected while minimizing false negatives.

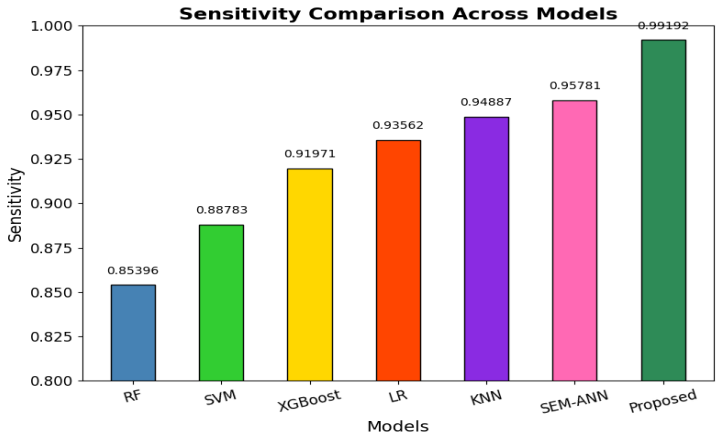

Figure 5: Sensitivity analysis

The F-measure or F1-score represents a balance of two metrics that is the harmonic mean of precision and recall. All of them performed decently, with RFs (85.40%), SVMs (88.78%), and XGBoost (91.97%). LRs (93.56%), KNNs (94.89%), and SEM-ANNs (95.79%) could do much better in this balance. The proposed model was better than all others, scoring an F1 of 99.20%, proving efficient in handling class imbalance while maintaining an excellent classification performance.

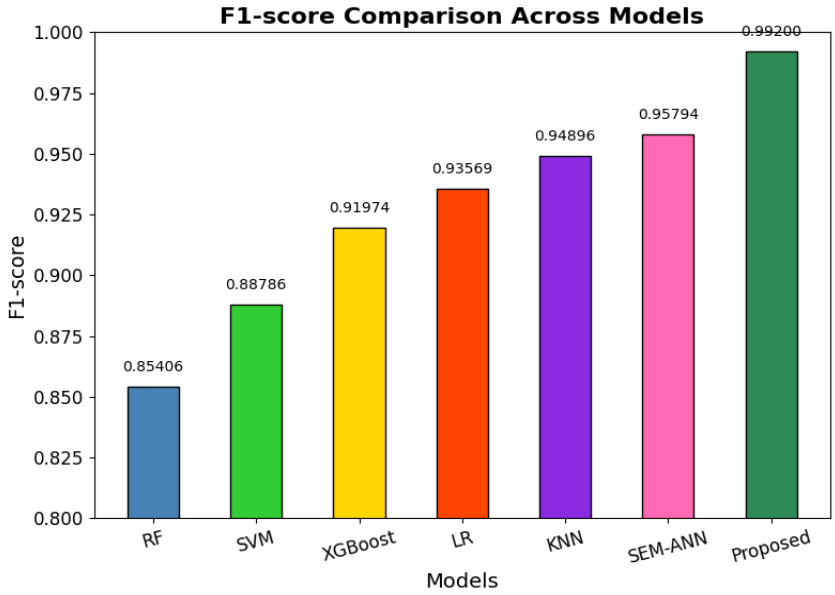

Figure 6: F1-score analysis

Specificity refers to the correct indication of normal traffic without giving false indications as an alarm. All three models provide arguably decent performance: RF (85.39%), SVM (88.78%), and XGBoost (91.97%). KNN (94.88%) and SEM-ANN (95.77%) turned the data even more specific. This proposed model achieved the highest specificity, i.e., 99.18%, which also implies the least misclassification of normal traffic as attacks.

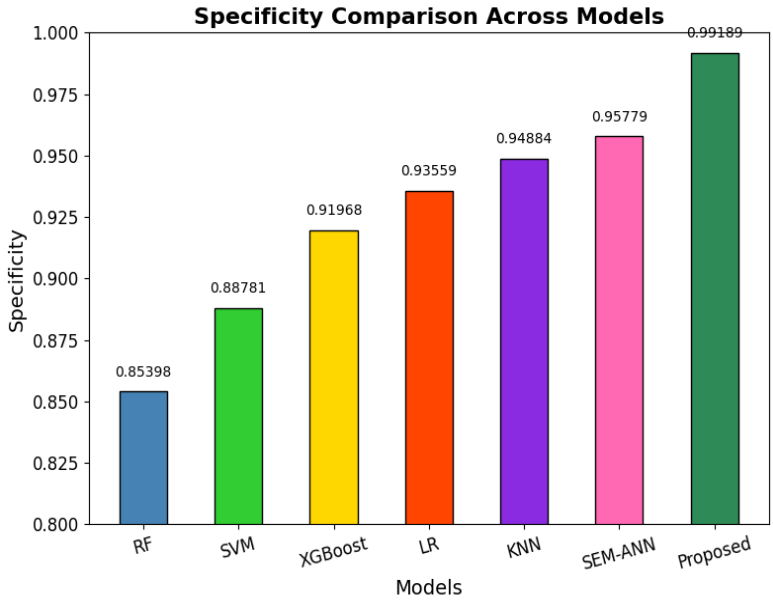

Figure 7: Specificity analysis

FPR defines how much normal traffic is wrong as an attack. The RF shows higher false positives with a higher percentage (4.27%), followed by SVM (3.94), and XGBoost (3.51). However, when introduced to LR (3.36%), KNN (3.06%), and SEM-ANN (2.95%), the false alarms significantly dropped. The suggested model exhibited the lowest FPR (2.49%), resulting in fewer false detections of typical traffic as malicious.

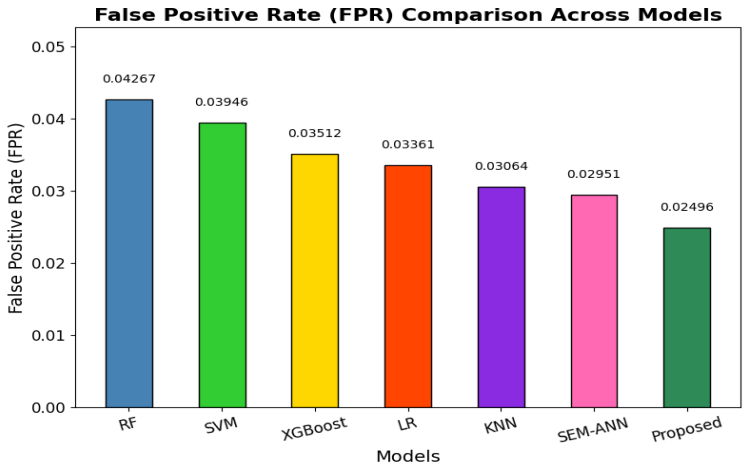

Figure 8: FPR analysis

FNR refers to how many real attacks were not discovered by the model. The false negatives are moderate for RF (4.95%), SVM (4.61%), and XGBoost (4.36%), while LR (4.25%) and KNN (3.98%) have shown an improved performance. In addition to that, SEM-ANN (3.83%) minimized false negatives. Thus, the proposed model has the least FNR (3.54%), thereby declaring its efficiency in recognizing most attack instances with a minimum of false negatives.

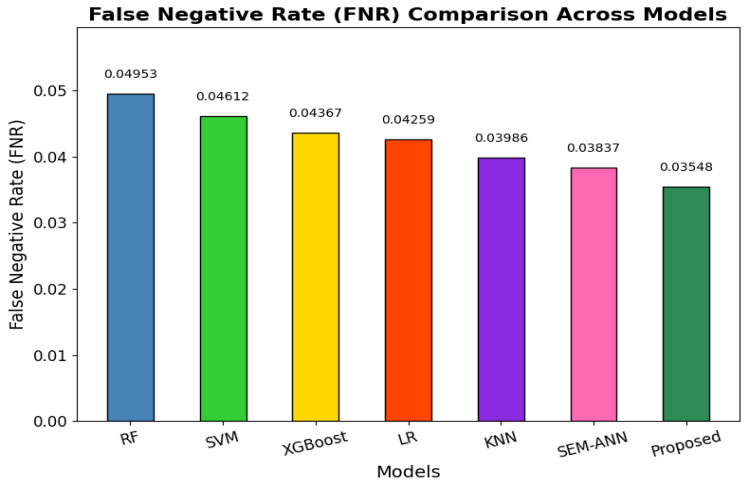

Figure 9: FNR analysis

Indeed, the model surpassed all other traditional models of ML significantly in all essential performance metrics. It recorded its supremacy in all aspects, that is, maximum accuracy (99.13%), maximum precision (99.13%), maximum sensitivity (99.19%), and minimum rates of false positives (2.49%) and false negatives (3.54%). High reliability in DDoS attack detection has been proven by the results. It was very much informed by the performed analyses that the required status performance of an IDS is achieved using combined feature selection, hyperparameter optimization, and state-of-the-art ML techniques.

5.2 Discussion

Below is the ROC curve representing the value for the design. The Receiver Operating Characteristic (ROC) curve is a major tool by which binary classification models are evaluated when there is TPR against FPR trade-off at varying thresholds of classification. The "Proposed Model" is the blue curve in the ROC graph, which displays how it performs at different thresholds. Superior performing models will have curves that sharply rise towards the top-left corner, indicating high true positive rates relative to low false positive rates. The area under the curve (AUC) gives an overall measure of model performance and has a value of 0.96, indicating strong discrimination capabilities in the model between normal and attack traffic. Such observations signify that the proposed model effectively detects attacks, while false positives remain at very low levels, and thus it is a good fit for intrusion detection applications. The ROC curve of the designed technique is shown in Fig. 10.

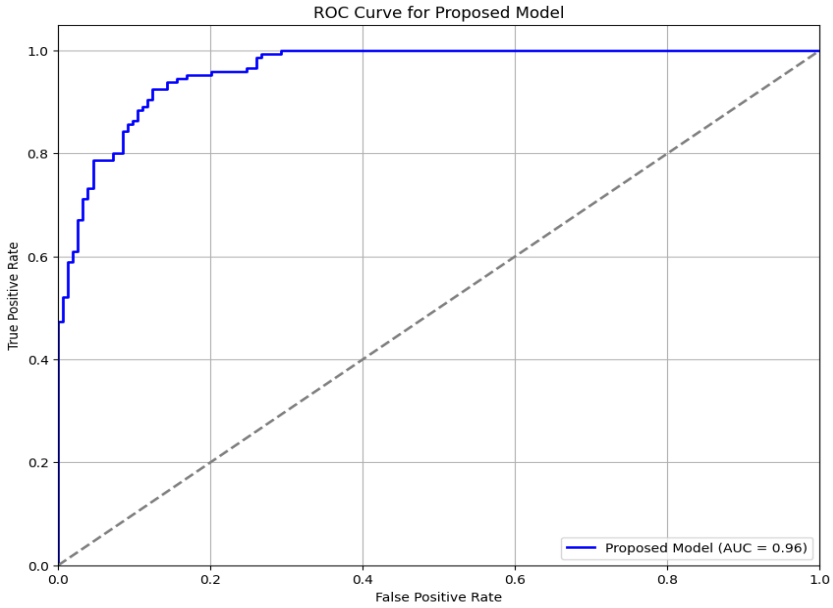

Figure 10: ROC curve

The picture here explains two graphs: the performance of a ML model during training and validation concerning accuracy and trends of loss values across multiple epochs. The left graph portrays training and validation accuracy at different epochs with an initial increase indicating the effective learning of a model after granting access to the dataset. The right graph illustrates loss figures against the different epochs, which otherwise indicate a reciprocal measure of accuracy. At the beginning, the training and validation losses both confirming the model's ability to minimize errors. Fig.11 illustrate the proposed model accuracy and loss performance.

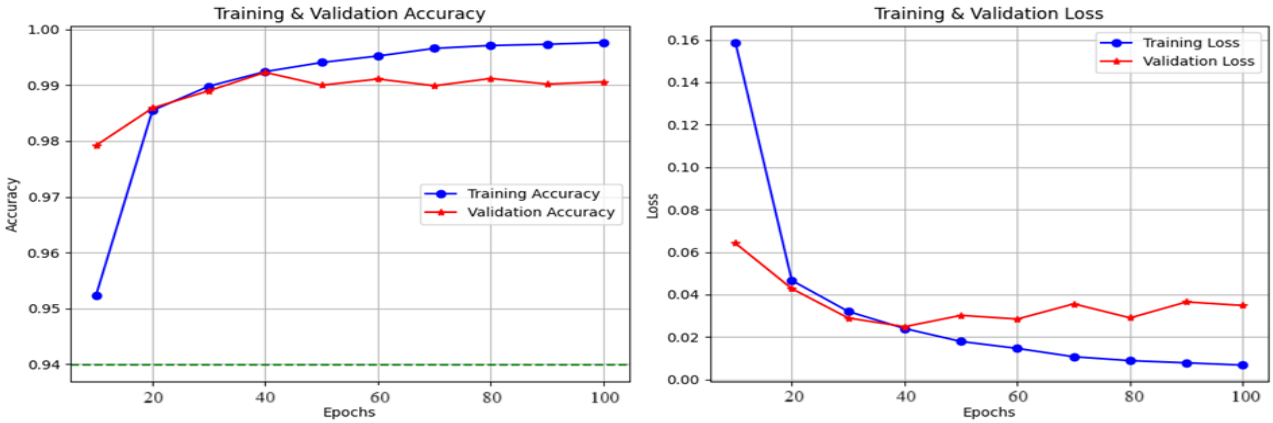

Figure 11: Accuracy and loss graph

The proposed approach for intrusion detection is based on the CSE-CIC-IDS2018 dataset. The methodology was implemented in Python with the use of software libraries such as TensorFlow and Scikit-learn, and tested against established ML models according to the performance metrics of accuracy, precision, sensitivity, F1-score, specificity, FPR, and FNR. The results indicate that the proposed method outperformed all other methods along these metric lines. Noteworthy is that it did achieve maximum accuracy (99.13%), precision (99.13%), and sensitivity (99.19%), alongside the lowest incidence of FPR (2.49%) and FNR (3.54%). Enhanced performance has been attributed to feature selection along with hyperparameter optimization and advanced ML techniques by the authors. Also, below are the ROC curve and the AUC (0.96) of the proposed model as proof of its strong discriminative ability. Thus, all together, the results very strongly support improved intrusion detection needs with the method; however, further work is necessary to overcome such overfitting.

6. CONCLUSION

Digitalization integration into healthcare has transformed the industry on scales where most

benefits lie within diagnoses, remote monitoring activities, and personalized treatment. All these changes are promises from information technology (IT), the IoT, and AI, thus providing richer data necessary for improving the effectiveness and quality of in-service delivery systems. However, such an event continues to expose a lot of cybersecurity threats and risks while relying on digital platforms, thus setting up the need for intelligent threat detection and mitigation strategies. HERFC-SWO is quite promising as it improves classification accuracy while providing an additional security layer through an AMFA protocol, ensuring the safe transmission of data across cloud networks. This comprehensive approach that marries advanced ML and good authentication measures will ensure the continued integrity and confidentiality of healthcare systems from evolving cyber threats. The proposed model outperformed all conventional ML models across all key performance metrics. It achieved the highest accuracy (99.13%), highest precision (99.13%), highest sensitivity (99.19%), and lowest false positive (2.49%) and false negative (3.54%). Future work can investigate federated learning for privacy-preserving threat detection within healthcare networks, thereby minimizing exposure of data while having models functioning optimally.

Compliance with Ethical Standards

Conflict of interest

The authors declare that they have no conflict of interest.

Human and Animal Rights

This article does not contain any studies with human or animal subjects performed by any of the authors.

Informed Consent

Informed consent does not apply as this was a retrospective review with no identifying patient information.

Funding: Not applicable

Conflicts of interest Statement: Not applicable

Consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material:

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Code availability: Not applicable

Competing Interests: Not applicable

REFERENCES

Silvestri, S., Islam, S., Amelin, D., Weiler, G., Papastergiou, S., & Ciampi, M. (2024). Cyber threat assessment and management for securing healthcare ecosystems using natural language processing. *International Journal of Information Security*, *23*(1), 31-50.

Tariq, M. U. (2024). Enhancing cybersecurity protocols in modern healthcare systems: Strategies and best practices. In *Transformative Approaches to Patient Literacy and Healthcare Innovation* (pp. 223-241). IGI Global.

Rajak, A., & Tripathi, R. (2024). DL-SkLSTM approach for cyber security threats detection in 5G enabled IIoT. *International Journal of Information Technology*, *16*(1), 13-20.

Admass, W. S., Munaye, Y. Y., & Diro, A. A. (2024). Cyber security: State of the art, challenges and future directions. *Cyber Security and Applications*, *2*, 100031.

George, A. S., Baskar, T., & Srikaanth, P. B. (2024). Cyber threats to critical infrastructure: assessing vulnerabilities across key sectors. *Partners Universal International Innovation Journal*, *2*(1), 51-75.

Ravi, V. (2024). Deep learning-based network intrusion detection in smart healthcare enterprise systems. *Multimedia Tools and Applications*, *83*(13), 39097-39115.

Alzubi, J. A., Alzubi, O. A., Qiqieh, I., & Singh, A. (2024). A blended deep learning intrusion detection framework for consumable edge-centric iomt industry. *IEEE Transactions on Consumer Electronics*.

Khan, I. A., Razzak, I., Pi, D., Khan, N., Hussain, Y., Li, B., & Kousar, T. (2024). Fed-inforce-fusion: A federated reinforcement-based fusion model for security and privacy protection of IoMT networks against cyber-attacks. *Information Fusion*, *101*, 102002.

Abidi, M. H., Alkhalefah, H., & Aboudaif, M. K. (2024). Enhancing Healthcare Data Security and Disease Detection Using Crossover-Based Multilayer Perceptron in Smart Healthcare Systems. *CMES-Computer Modeling in Engineering & Sciences*, *139*(1).

Adil, M., Khan, M. K., Kumar, N., Attique, M., Farouk, A., Guizani, M., & Jin, Z. (2024). Healthcare internet of things: Security threats, challenges and future research directions. *IEEE Internet of Things Journal*.

Vegesna, V. V. (2024). Machine Learning Approaches for Anomaly Detection in Cyber-Physical Systems: A Case Study in Critical Infrastructure Protection. *International Journal of Machine Learning and Artificial Intelligence*, *5*(5), 1-13.

Sun, Z., An, G., Yang, Y., & Liu, Y. (2024). Optimized machine learning enabled intrusion detection 2 system for internet of medical things. *Franklin Open*, *6*, 100056.

Isakov, A., Urozov, F., Abduzhapporov, S., & Isokova, M. (2024). Enhancing Cybersecurity: Protecting Data In The Digital Age. *Innovations in Science and Technologies*, *1*(1), 40-49.

Nanjappan, M., Pradeep, K., Natesan, G., Samydurai, A., & Premalatha, G. (2024). DeepLG SecNet: Utilizing deep LSTM and GRU with secure network for enhanced intrusion detection in IoT environments. *Cluster Computing*, 1-13.

Prince, N. U., Faheem, M. A., Khan, O. U., Hossain, K., Alkhayyat, A., Hamdache, A., & Elmouki, I. (2024). AI-powered data-driven cybersecurity techniques: Boosting threat identification and reaction. *Nanotechnology Perceptions*, *20*, 332-353.

Ala'a, M., Ramayah, T., & Al-Sharafi, M. A. (2024). Exploring the impact of cybersecurity on using electronic health records and their performance among healthcare professionals: A multi-analytical SEM-ANN approach. *Technology in Society*, *77*, 102592.

Ezekwueme, A. E., Abel, C. E., & Dike, G. A. (2024). Efficiency of IoT Adoption and Supply Chain Optimization: An Empirical Evidence from Nigeria.

Nagamalla, V., & Sanapala, R. K. (2023). Integrating Predictive Big Data Analytics with Behavioral Machine Learning Models for Proactive Threat Intelligence in Industrial IoT Cybersecurity. *International Journal of Wireless & Ad Hoc Communication*, *7*(2).

Radanliev, P., & De Roure, D. (2022). Advancing the cybersecurity of the healthcare system with self-optimising and self-adaptative artificial intelligence (part 2). *Health and Technology*, *12*(5), 923-929.

Rajamäki, J., Rathod, P., Ferreira, J. C., Ahonen, O., Serrão, C., & do Carmo Gomes, M. (2024, May). Enhancing Cybersecurity Education for the Healthcare Sector: Fostering Interdisciplinary ManagiDiTH Approach. In *2024 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-7). IEEE.

Nadhan, A. S., & Jacob, I. J. (2024). Enhancing healthcare security in the digital era: Safeguarding medical images with lightweight cryptographic techniques in IoT healthcare applications. *Biomedical Signal Processing and Control*, *88*, 105511.

Arunprasath, S., & Annamalai, S. (2024). Improving patient centric data retrieval and cyber security in healthcare: privacy preserving solutions for a secure future. *Multimedia Tools and Applications*, 1-31.

Elnawawy, M., Hallajiyan, M., Mitra, G., Iqbal, S., & Pattabiraman, K. (2024). Systematically Assessing the Security Risks of AI/ML-enabled Connected Healthcare Systems. *arXiv preprint arXiv:2401.17136*.

Jimmy, F. N. U. (2024). Cyber security Vulnerabilities and Remediation Through Cloud Security Tools. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023*, *2*(1), 129-171. https://www.kaggle.com/datasets/solarmainframe/ids-intrusion-csv

Chrenka, E. A., Dehmer, S. P., Maciosek, M. V., Essien, I. J., & Westgard, B. C. (2024). Use of Sequential Hot-Deck Imputation for Missing Health Care Systems Data for Population Health Research. *Medical Care*, *62*(5), 319-325.

Qu, K., Xu, J., Hou, Q., Qu, K., & Sun, Y. (2023). Feature selection using Information Gain and decision information in neighborhood decision system. *Applied Soft Computing*, *136*, 110100.

Beattie, J. R., & Esmonde-White, F. W. (2021). Exploration of principal component analysis: deriving principal component analysis visually using spectra. *Applied Spectroscopy*, *75*(4), 361-375.

Trizoglou, P., Liu, X., & Lin, Z. (2021). Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines. *Renewable Energy*, *179*, 945-962.

Alduailij, M., Khan, Q. W., Tahir, M., Sardaraz, M., Alduailij, M., & Malik, F. (2022). Machine-learning-based DDoS attack detection using mutual information and random forest feature importance method. *Symmetry*, *14*(6), 1095.

Arunkumar, M., & Kumar, K. A. (2023). GOSVM: Gannet optimization based support vector machine for malicious attack detection in cloud environment. *International Journal of Information Technology*, *15*(3), 1653-1660.

Abdel-Basset, M., Mohamed, R., Jameel, M., & Abouhawwash, M. (2023). Spider wasp optimizer: a novel meta-heuristic optimization algorithm. *Artificial Intelligence Review*, *56*(10), 11675-11738.