INTERNATIONAL JOURNAL OF Telecommunications
AIR DEFENSE COLLEGE

IoT-Based Cybersecurity Threat Detection Using Feature Selection and Dipper Throated Optimization Algorithm

Ali E. Takieldeen[1], S.K.Towfek[2,3],Marwa Metwally[1,4], Ahmed M. Zaki[1], Nima Khodadadi[5]

1 Faculty of Artificial Intelligence,Delta University for Science & Technology,Mansoura, Egypt
2 Computer Science and Intelligent Systems Research Center, Blacksburg 24060, Virginia, USA
3 Applied Science Research Center, Applied Science Private University, Amman, Jordan
4 Jadara University Research Center, Jadara University, Jordan
5 Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL, USA
Emails: a_takieldeen@deltauniv.edu.eg,sktowfek@jcsis.org,mmm@ieee.org, azaki@jcsis.org, nima.khodadadi@miami.edu
Corresponding author(s): S.K.Towfek (sktowfek@jcsis.org)

Citation:
Takieldeen , A.; Towfek , S.; Metwally, M.; Zaki, A., Khodadadi, N.
Inter. Jour. of Telecommunications, ***IJT'2025,***
Vol. 05, Issue 02, pp. 1-22, 2025.

Doi: 10.21608/ijt.2025.387473.1106

Editor-in-Chief: Youssef Fayed.

Received: 21/05/2025.

Accepted: date 20/07/2025.

Published: date 20/07/2025.

Publisher's Note: The International Journal of Telecommunications, IJT, stays neutral regard-ing jurisdictional claims in published maps and institutional affiliations.

Abstract: Technology on the ascent has no less disturbing our lives as the progression in digital platforms has not slumped the importance of developing solid security key measures to protect a computer system or network from unscrupulous individuals. This work aims to investigate enhancing programs employed in threat identification associated with cybersecurity in the background of the IoT in combination with the Dipper Throated Optimization (DTO) algorithm and Gradient Boosting. The increasing intricacy of information systems and a sharp increase in the usage of IoT devices would indicate that technology's need to prevent information leakage or data breaches is becoming more critical. Facilitating the management of the arising challenges related to optimization problems in the field of cybersecurity is the playing ground of metaheuristic optimization algorithms based on the principles of natural sciences. These algorithms are well described in the literature, and this research carefully analyzes and deploys them to carry out feature selection with a focus on the IoT cybersecurity context. Specifically, they solve the typically tricky combinatorial optimization problem of binary optimization to feature selection to pick the most relevant features with the most negligible computation intelligently. Another increase in efficiency when applying the cybersecurity framework is when it is integrated with machine learning models. For the regression, the following approaches have been implemented: Gradient Boosting, CatBoost, and XGBoost. Besides, mean squared error (MSE) and the percentage of change in root mean squared error (RMSE) were used when comparing these models. The results of this research advance the scholarship of optimization in the context of IoT cybersecurity and hold practical implications for improved threat detection models' implementation in applications. Including DTO with Gradient Boosting enhances the attainment of high-quality cybersecurity threat detection in IoT, ensuring the value of speeding up modified interconnected systems.

Keywords: Optimization; IoT cybersecurity attacks; Dipper Throated Optimization algorithm; Gradient Boosting; accuracy; feature selection

1. Introduction

The defense of computer systems and networks from malicious actors is one prominent feature in the complex world of recent technology. Cybersecurity is critical in preventing unauthorized disclosures of information and theft of hardware and software [1]. The multi-layered nature of information systems, coupled with the widespread adoption and rise in smart devices and technology, including the Internet of Things (IoT) [2], increases the difficulties encountered in this domain. Cybersecurity takes on the significance of ensuring that the major systems like power distribution, electoral processes and financial systems are secure because they have more significant implications beyond their logical response to physical reality [3]. There is an aspect beyond the world of cybersecurity where using artificial intelligence (AI) and optimization techniques become helpful in dealing with these problems that cross over [4-5]. Optimization is ubiquitous in all sectors, from engineering design to economics and even holiday trip planning [6]. It is essential to properly manage several limited resources, whether monetary, time or any other types of assets, to generate optimal results [7-8]. Typically, real-world optimization scenarios involve nonlinearity, multimodality, and intricate constraints with often conflicting objectives; real-world optimization situations readily optimize this factor to key indications while overlooking others. Solving an optimal or sub-optimal problem is quite challenging and emphasizes the need for thinking outside the box.

The ideal innovative solution is using metaheuristic optimization algorithms, each inspired by a specific biological or natural effect/phenomenon [9-10]. The Sine Cosine Algorithm (SCA), based on the periodic features of sine and cosine functions, theory utilizing their periodicity is used when exploring and exploiting the search space when it comes to finding pockets between a good solution and Zero value. Natural evolution, inspired by the navigation method of moths that utilizes a sort of transverse orientation to optimize solutions, is used with Moth-Flame Optimization (MFO) [11-12]. Alternatively, Particle Swarm Optimization adjusts candidate solutions step-by-step by imitating particle movement in a search space [13 - 14].The Whale Optimization Algorithm (WOA) mimics the hunting practice of humpback whales in general and its bubble-net strategy for them [15]. The Grey Wolf Optimizer replicates grey wolves' leadership structure and hunting behaviors and uses alpha, beta, delta, and omega wolves for simulation [16-17]. Finally, the Firefly Algorithm (FA) is based on the social behavior of fireflies and offers a distinct optimization approach [18]. Binary optimization for feature selection is a significant technique within the scope of cybersecurity [19]. This approach involves choosing the most representative features from a dataset, which is inherently problematic for combinatorial optimization [20]. Therefore, binary optimization algorithms, designed to work with binary values that can either be 0 or 1, are central in selecting a subset of features that best contribute towards the optimal performance of a model, thereby helping reduce complexity during computation [21].

Furthermore, machine learning becomes critical for integration into this framework. A subset of AI is machine learning, which entails machines' ability to replicate intelligent human behavior, solving intricate issues in a way that humans do [22]. These models have included machine learning models, including Gradient Boosting, CatBoost, XGBoost, Linear Regression and Multilayer Perceptron, Random Forest Decision Tree K Nearest Neighbours KNN, Extra Trees As a collection of regressive tasks in the new standard patterns of the cybersecurity realm [23-25].

This introduction paves the way for a detailed investigation into the complex relationship between AI, optimization algorithms, feature selection and machine learning models to improve the efficacy of cybersecurity systems. They subsequently discuss the intricacies of related works, materials met, hods, and experimental results before concluding the course while identifying avenues for future research.

The main contributions of this study can be summarized in:

1. This study significantly contributes to data analysis and investigation, offering a detailed exploration of a Kaggle dataset, including malware attacks and visualization through geographical data and scatter plots.

2. This research deals with feature selection and uses miscellaneous binary optimization algorithms, including bDTO, bSCA, bMFO, bPSO, bWAO, bGWO, and bFA for comparison.

3. Fundamental machine learning models, including Gradient Boosting, CatBoost, XGBoost, linear regression, Multilayer Perceptron, Random Forest, Decision Tree, K-nearest neighbors, and Extra Trees, serve as reference models for the evaluation.

4. The study discusses specific aspects, including brief reviews of specific optimization algorithms, such as the Sine Cosine Algorithm, Moth-Flame Algorithm, Particle Swarm Algorithm, Whale Optimization Algorithm, Grey Wolf Optimizer and Firefly Algorithm, as well as the mechanisms and inspiration behind these algorithms.

5. Hyperparameter tuning is performed using Dipper Throated Optimization to improve the effectiveness of the GB algorithm in general.

6. ANOVA and Wilcoxon Tests for Statistics determine performance evaluations and significance tests.

This manuscript's contributions offer expertise and a relevant literature review of feature selection, ML, and optimization algorithms, as well as statistical analysis of cybersecurity and optimization.

However, while this study proposes a promising approach through the integration of the Dipper Throated Optimization (DTO) algorithm with Gradient Boosting for IoT-based cybersecurity threat detection, certain limitations should be acknowledged. The evaluation relies on a single dataset, which may not fully represent the complexity and diversity of real-world IoT environments, potentially affecting the model's generalizability. Moreover, the computational demands of metaheuristic optimization, along with the risk of overfitting due to extensive hyperparameter tuning, may constrain the deployment of the proposed framework in resource-limited or real-time applications. These limitations suggest the need for future studies to validate the approach across varied datasets, test real-time performance, and explore its adaptability to emerging or adversarial threats.

The novelty of this work lies in the dual application of the Dipper Throated Optimization (DTO) algorithm—used in its binary form (bDTO) for feature selection and in its continuous form for hyperparameter optimization of the Gradient Boosting model. While prior research has employed metaheuristic algorithms in isolation, this study uniquely integrates bDTO with a full machine learning pipeline that includes extensive model benchmarking and statistical validation using ANOVA and Wilcoxon tests. Moreover, the proposed framework demonstrates superior performance compared to several state-of-the-art methods, including CatBoost, XGBoost, and other binary optimizers such as bPSO and bGWO. Unlike existing approaches, which often neglect either feature selection or model tuning, our methodology addresses both, resulting in a more efficient, interpretable, and accurate IoT cybersecurity threat detection system.

2 Related Works

In cybersecurity, several researchers depend on excellent and reliable intrusion detection systems for IoT-associated devices and distribute heterogeneous devices. A recent work [26] has illustrated the heightened need to have a high-speed and reliable intrusion detection system for IoT appliances and bodies that are dispersed and distinct. Based on the ideas of Decision Trees, this paper presents an intelligent intrusion detection model and carefully thinks about how to rank security features. Practical application of the model on an accurate data set for network intrusion detection systems demonstrates the model's effectiveness. The model's ability to predict and find cyberattacks with less computational complexity than traditional machine learning methods is supported by performance evaluation metrics like accuracy, precision, recall, and F-score. Moving our attention to the smart grid domain [27] emphasizes how vital this cutting-edge power system is to have accurate intrusion detection and response systems. Together, a whale optimization algorithm (WOA) and an artificial neural network (ANN) are used in this paper to create a new intrusion detection model. Getting the ANN's weight vector to have the most minor mean square error depends on the WOA. This cutting-edge WOA ANN model classifies binary, triple, and multi-class cyberattacks and power system incidents. The WOA-ANN model is better than other commonly used classifiers at dealing with the complex

problems of attacks, failure prediction, and failure detection in the smart grid. This was proven by using databases of power-system attacks as proof.

Artificial intelligence (AI) is a crucial technology for protecting Internet-connected systems from cyber threats due to the Fourth Industrial Revolution, which was discussed in [28]. AI has many uses, and this paper looks at some of them. These include machine learning, deep learning, natural language processing, and rule-based expert systems modeling. The idea of "AI-driven Cybersecurity" emerges by utilizing these AI techniques, promising automated and intelligent solutions that are better than traditional security systems. Insights into intelligent computing and AI-based approaches to cybersecurity challenges focus on delivering a comprehensive guide for cybersecurity researchers and industry professionals. The transformative role of AI in cybersecurity is further explored in [29], which adds nuance to this story. The fact that it cuts through the daily security alerts shows that AI can provide instant insights. One of the main parts of the study is visual analysis, which looks at how AI can be used in cybersecurity. There is a close look at how AI has changed the structure of cybersecurity law. The paper talks about essential research areas that could significantly impact how AI security systems are developed in the future. Examples include face recognition and deep neural networks for speech recognition. It uses five evaluation factors and a heat map to show the global landscape of AI applications in cybersecurity research in a way that has never been seen before. This study is a valuable tool for planning and making decisions in the field.

Using machine learning algorithms in cybersecurity is hard because they can be fooled by adversarial examples [30]. This paper introduces a novel approach called the brute force attack method to evaluate the robustness of machine learning classifiers against negative examples in cybersecurity. By being straightforward and practical, this method, which operates in a black-box manner, addresses flaws in current harmful attack methods. Among the security systems being tested are host intrusion detection, Android malware detection, and network intrusion detection systems that use machine learning. Starting with early results, the suggested method seems faster at computing and better at protecting against attacks using cutting-edge generative adversarial networks. This makes it a valuable tool for testing different machine learning systems' safety in cybersecurity. The paper then discusses the escalating concern of finding cybersecurity attacks and cyber anomalies [31]. The authors present "CyberLearning," a thorough security modeling framework utilizing artificial intelligence and machine learning techniques. This framework uses correlated feature selection and an empirical analysis to find out how well different machine learning-based security models work. The study evaluates the effectiveness of ten other classification techniques and an artificial neural network-based model for anomaly detection and binary classification for various cyberattacks. These papers use well-known security datasets, UNSW-NB15 and NSL-KDD, to try to become a valuable resource for data-driven security modeling. They hope to contribute to the changing world of cybersecurity with their insights and findings.

The security of modern intelligent power grids is a problematic problem addressed in [32] in detailed paper. It is suggested that an attack detection model based on machine learning be used because both natural and artificial events could mess up power systems. In addition to feature construction engineering, the model uses data and logs gathered by phasor measurement units (PMUs). As the primary classifier for AdaBoost, Random Forest is selected. The evaluation of open source simulated power system data with 37 event scenarios shows how well the model works. It beats more than eight new techniques with an accuracy rate of 93.91% and a detection rate of 93.6%. This reveals the ability of machine learning to make power systems safer. In the context of dynamic data-driven vulnerability assessments, [33] introduces a cognitive cybersecurity approach to deal with the problems brought on by large amounts of different types of data in Security Operations Centres (SOCs). Because security repositories are often full of various kinds of information and duplicates, manual vulnerability assessment methods often produce wrong data. Getting rid of conflicting vulnerability reports and preprocessing embedded security indicators are two ideas in the paper that can help you make reliable datasets. An ensemble meta-classifier method is introduced to combine different machine learning techniques to improve predictive accuracy over single algorithms. The proposed cognitive security methodology shows promise by better addressing incompleteness and diversity across cybersecurity alert repositories. The experimental analysis of actual cybersecurity data sources provides insights into the ability of the selective ensemble methodology to infer patterns of computer system vulnerabilities.

Table 1 offers an overview of previous works in cybersecurity, paying particular attention to threat detection in IoT and other distributed systems. These studies propose machine learning, artificial intelligence, metaheuristic optimization algorithms, and advanced methods to improve cybersecurity solutions. From proposing new and efficient intrusion detection systems for IoTs within a short time to developing specific models for smart grid security, all these works have highlighted different methodologies and their efficiency in handling security threats. AI and machine learning have been vital in enhancing automated and intelligent systems, hence enhancing the detection and prevention of cyber threats. In this way, each of the references plays a distinctive part in developing new approaches to cybersecurity and proposes ideas and recommendations for further studies and real-life implementations.

These studies show why advanced technologies, such as machine learning and artificial intelligence (AI), are increasingly crucial for improving cybersecurity. New intrusion detection models, practical methods for assessing cyber-attacks, and unified security architectures are all presented in these works. Every contribution adds something valuable to the conversation about how cybersecurity constantly changes. As these studies show, Internet of Things (IoT) devices can be vulnerable differently. On the other hand, they also demonstrate how security solutions for smart grids can be improved. Cybersecurity is getting harder to handle because computers are becoming more complicated and weak. To overcome this problem, we need technologies that work well with each other and are advanced enough to adapt to new issues that come up with different computer parts.

Table 1: Summary of Related Works in Cybersecurity and IoT

Ref	Focus Area	Key Contributions	Methods/Algorithms Used	Results	publication date	Limitation
[26]	IoT Intrusion Detection	Development of fast and reliable intrusion detection systems for IoT devices	Decision Trees	Demonstrated effectiveness in predicting and finding cyberattacks with lower computational complexity	2021	No optimization or feature selection techniques used
[27]	Smart Grid Intrusion Detection	Creation of an intrusion detection model for smart grids	Whale Optimization Algorithm (WOA), Artificial Neural Network (ANN)	Effective in classifying binary, triple, and multi-class cyberattacks and power system incidents	2020	Limited to smart grid context; no comparative model testing
[28]	AI in Cybersecurity	Examination of AI-driven cybersecurity approaches	Machine Learning, Deep Learning, Natural Language Processing, Rule-based Expert Systems	Highlighted AI techniques for automated and intelligent cybersecurity solutions	2021	Theoretical; lacks empirical evaluation or optimization focus

[29]	AI Applications in Cybersecurity	Exploration of AI's transformative role in cybersecurity	Visual Analysis, AI Law	Provided insights into AI's impact on cybersecurity law and instant insights from security alerts	2019	Focused on legal and policy context; lacks technical solution
[30]	Adversarial Machine Learning	Evaluation of machine learning classifiers' robustness against adversarial examples	Brute Force Attack Method	Showed improved protection against attacks using generative adversarial networks	2020	No feature selection or metaheuristic application
[31]	Cyber Attack Detection	Development of a security modeling framework using AI and machine learning	Cyber-Learning, Correlated Feature Selection, Empirical Analysis	Evaluated the effectiveness of classification techniques and ANN-based models for anomaly detection	2021	Generic models; no optimization-enhanced approaches
[32]	Intelligent Power Grids	Machine learning-based attack detection model for power grids	AdaBoost, Random Forest, Phasor Measurement Units (PMUs)	Achieved high accuracy (93.91%) and detection rate (93.6%) in power system security	2019	Domain-specific; not applicable to broader IoT settings
[33]	Cognitive Cybersecurity	Cognitive approach to vulnerability assessment in SOCs	Ensemble Meta-Classifier, Data Preprocessing	Improved predictive accuracy and addressed issues of data incompleteness and diversity in security alerts	2021	Does not address IoT-specific architecture or constraints

3. Materials and Methods

The following section gives details on the research design used in this study. The study is based on an extensive dataset solely targeted to cybersecurity threat identification; this dataset is used to predict and test the different machine learning solutions and the Dipper Throated Optimization (DTO) algorithm. Information related to the dataset and DTO algorithm, types of ML models, and gradient-boosting hyperparameter optimization is also discussed.

To clarify the integration of components in our framework, a methodological flowchart is introduced (see Figure 1. The process begins with the acquisition and preprocessing of the cybersecurity dataset, which involves normalization and the handling of missing or categorical values. Following this, binary feature selection is performed using the Binary Dipper Throated Optimization (bDTO) algorithm. This step reduces data dimensionality while retaining informative attributes relevant to cybersecurity threat detection.

Once the optimal feature subset is selected, the dataset is partitioned into training and testing sets. A suite of machine learning models—including Gradient Boosting, CatBoost, XGBoost, and others—is then trained using the selected features. Specifically, the Gradient Boosting model is further optimized through hyperparameter tuning with the original DTO algorithm to enhance predictive performance. Finally, model performance is evaluated using standard regression metrics such as MSE, RMSE, and R^2, along with statistical validation using ANOVA and Wilcoxon tests. This integrated pipeline ensures both high accuracy and statistical robustness in identifying cyber threats across IoT systems.

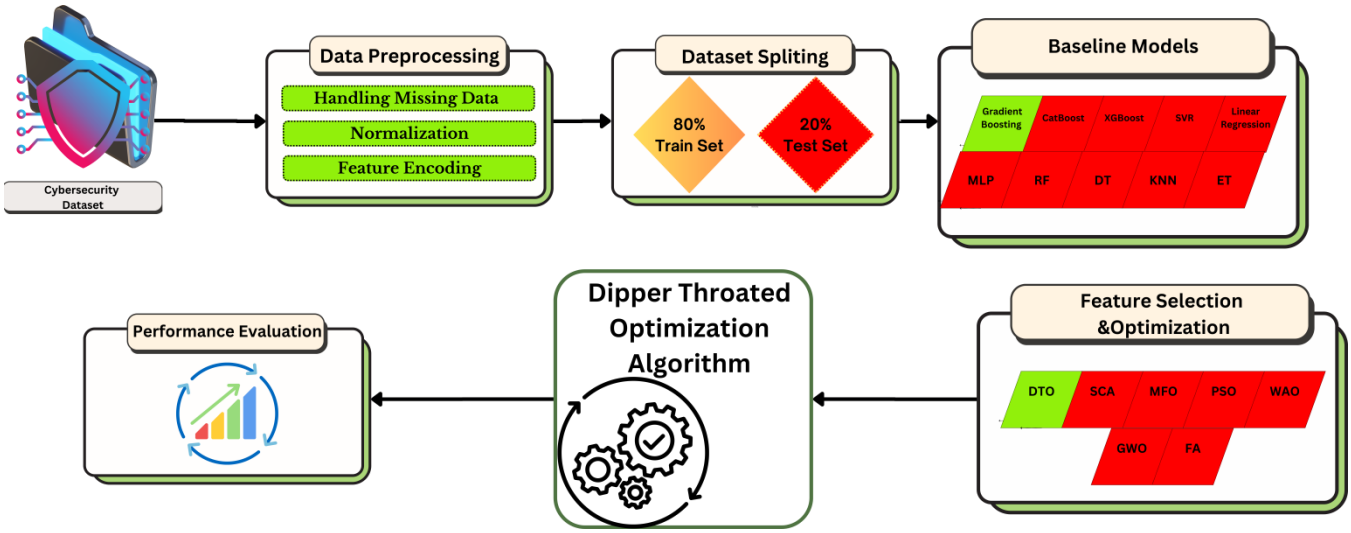

Figure 1: Flowchart of the proposed cybersecurity threat detection framework

3.1 Dataset

The prerequisite of our work is the accurate dataset collected from Kaggle [34] using the filtration process required for cybersecurity threat detection. This specific data set comprises a range of aspects that help comprehend and prevent possible cyber risks. The structure of analyzed dataset involves characteristics like Timestamp, Source IP Address, Destination IP Address, Source Port, Destination Port, Protocol, Packet Length, Packet Type, Traffic Type, Payload Data, Malware Indicators, Anomaly Scores, Alerts/Warnings, and Attack Type This range of features allows including the optimization algorithms and machine learning definitions into a comprehensive understanding of the research contributing the solidity and relevance of the results gained.

The data allows for the plotting of the distribution of malware attacks by region, as shown in Figure 2. This geographical distribution offers valuable insights that can aid in understanding the prevalence of cyber threats in specific regions with high attack rates. Recognizing the geographical aspect of these threats is therefore vital for devising accurate protective measures for cyberspace and enhancing the efficiency of threat identification mechanisms.

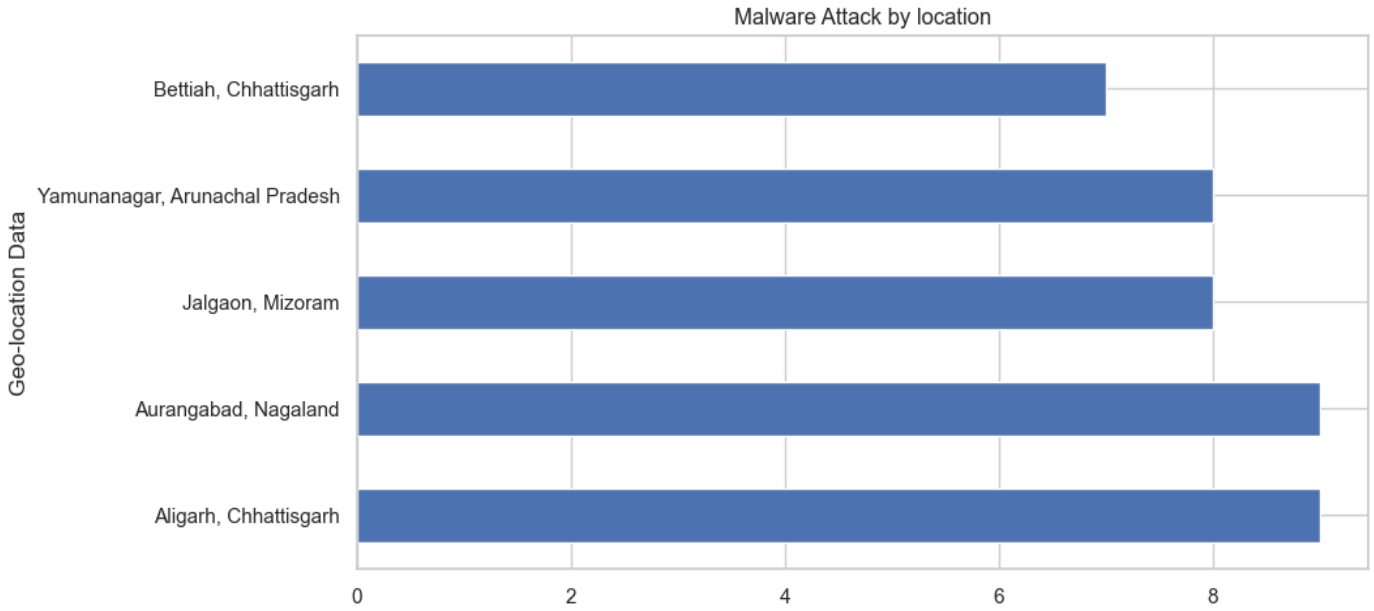

Figure 2: Malware Attack by location

Figure 3 presents the Scatter and Density Plot visualization of the distribution of packet lengths based on the attack type. This visualization aids in identifying patterns and establishing relationships between different attacks and their properties. By comparing the general distribution of packet lengths with other abnormal behaviors associated with specific cyber threats, researchers can effectively select the appropriate feature and model for subsequent stages of the study.

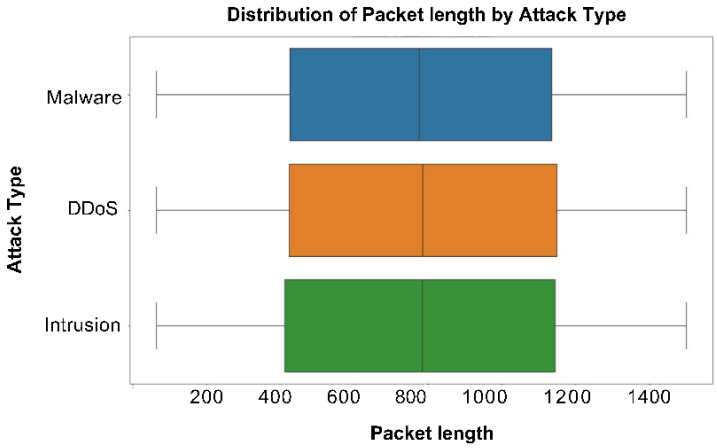

Figure 3: Distribution of Packet length by Attack Type

Figure 4 shows the quantity and frequency of cyber-attacks from 2020 to 2023. The temporal analysis yields information about the cyclical occurrence of threats and their frequency, which helps in understanding areas of higher risks and the efficacy of measures implemented. Such information is valuable for planning future preventive measures and distributing resources in the context of cybersecurity.

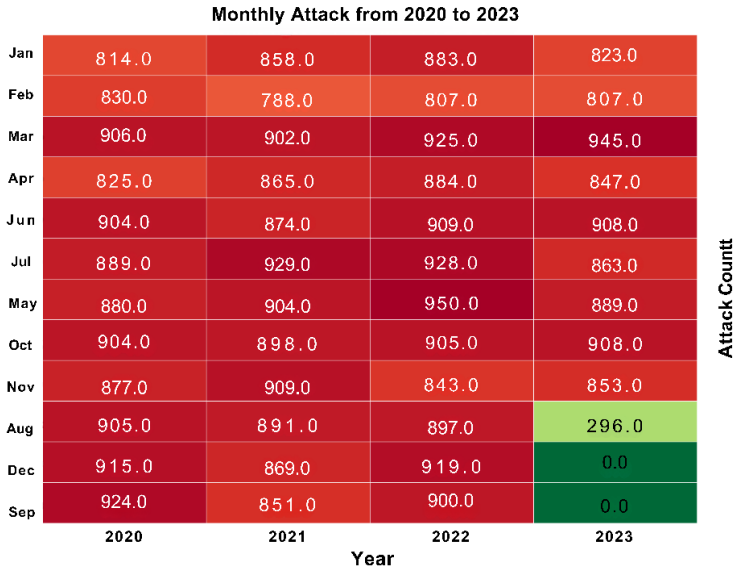

Figure 4: Monthly Attacks from 2020 to 2023

These visualizations above help us understand the dataset as a whole and provide insight into the different aspects of the data. Attributes significant for describing the structures and dependence necessary for constructing threat detection and prevention schemes are disclosed from geographical distributions, temporal trends, and data properties. It is then possible to identify regions or countries most vulnerable to the particular type of attack, study the activities and abnormalities linked to the various threats, and analyze patterns and oscillations of attacks over time. This examination is essential in data exploration to guarantee that the formulated cybersecurity plans are optimally suitable. These visualizations help enhance the threat detection models and promote proper and effective cybersecurity mechanisms.

3.2 Machine Learning Basic Models

Powerful algorithms form a multi-model machine learning approach to achieve the best performance while addressing the cybersecurity threat detection problem. Table 2 below demonstrates the type of machine learning models used in this study: All the models have unique skills for improving Cybersecurity Threats' detection efficiency and reliability. These models are incorporated with the DTO algorithm to enhance the threat detection systems.

Table 2: Machine Learning . Models

Model	Technique	Features	Benefits	Reference
Gradient Boosting	Boosting technique that builds sequential models	Corrects errors of predecessor models	Minimizes overall prediction error by setting target outcomes for each subsequent model	[35]
CatBoost	Gradient boosting variant	Handles categorical and numerical features seamlessly	Eliminates need for feature encoding techniques; introduces SWQS algorithm to handle missing values; reduces overfitting and enhances overall performance	[36]

Model	Technique	Features	Benefits	Reference
XGBoost	Gradient-boosting decision trees	Robust implementation	Famous for optimizing machine learning models	[37]
Linear Regression	Predictive analysis technique	Predicts the value of the dependent variable based on an independent variable	Fundamental for understanding relationships between variables	[38]
Multilayer Perceptron (MLP)	Versatile model used in various machine learning techniques	Adequate for classification and regression tasks	Known for providing highly accurate results in classification problems	[39]
Random Forest	Ensemble learning method	Constructs multiple decision trees during training	Output determined by majority class selected by most trees	[40]
Decision Tree	Non-parametric supervised learning algorithm	Used for both classification and regression tasks	Hierarchical structure with root nodes, branches, internal nodes, and leaf nodes	[41]
K-Nearest Neighbors (KNN)	Classification and regression technique	It relies on the principle that similar data points have similar labels or values.	Stores the entire training dataset as a reference during the training phase	[42]
Extra Trees	Ensemble supervised machine learning method	Utilizes extremely randomized trees	Employed for classification and regression tasks; contributes to overall model robustness	[43]

Altogether, these machine learning models are a versatile portfolio. Each method compensates for the other's weaknesses and possesses unique features that help to improve the cybersecurity threat detection process. The following sections will map these models to the DTO algorithm, where the effects on the enhancement of threat detection systems will be illustrated.

3.3 Dipper Throated Optimization Algorithm

The throated Optimization (DTO) algorithm [44-45] is one of the metaheuristic optimization techniques for solving complex optimization problems based on the navigation behaviors of the dipper and other foraging aquatic bird families having a particular throat. In our view, it forms a basis for our research on the methodology of developing more robust protection from cybersecurity threats. The DTO Algorithm described in

Algorithm 1 is devised to include tolerance and capability to solve a broad range of optimization problems related to cybersecurity threat identification.

Algorithm 1: Dipper Throated Optimization Algorithm

```
Initialization positions BPi (i = 1, 2, . . . , n) with size n,
velocities BVi (i = 1, 2, . . . , n), total number of iterations Tmax,
fitness function fn, c, C1, C2, C3, C4, C5, r1, r2, R, t = 1
Calculate objective function fn for each bird BPi
Find the best bird BPbest
    While t ≤ Tmax do
        for (i = 1: i < n + 1) do
            if (R < 0.5), then
                Update the position of the current swimming bird as
                 BPnd (t + 1) = BPbest(t) - C1.|C2.BPbest(t) - BPnd(t)|
             else
              Update the velocity of the current flying bird as
              BV (t + 1) = C3BV(t) + C4r2(BPbest(t) - BPnd(t)) + C5r2(BPGbest - BPnd(t))
              Update the position of the current flying bird as
              BPnd (t + 1) = BPnd(t) + BV(t + 1)
             end if
        end for
      Calculate objective function fn for each bird BPi
      Update c, C1, C2, R
      Find the best bird BPbest
      Set BPGbest = BPbest
      Set t = t + 1
Return the best bird BPGbest
```

In our research framework, the algorithm's flexibility provides a powerful tool for dealing with some very complex problems associated with threat detection in cybersecurity.

3.4 Gradient Boosting Hyperparameter Tuning

The Gradient Boosting algorithm entails fine-tuning several parameters to improve performance. These hyperparameters are rather important because they govern different facets of the learning process and the construction of the model as a whole; this influences the accuracy, speed, and generalization capability of the model on unseen data sets immensely. These parameters should, therefore, be fine-tuned to give the desired results of the algorithm suitable for the specific data sets and the problems encountered. Tuning hyperparameters avoid overtraining and increases the model generalization capacity and a better predictive capability.

Table 3 presents the values for the tree-specific parameters that were set to determine the gradient boost. These parameters are useful in managing overfitting and the proper management of split in the decision trees. Thus, they help shape a solid and not-depth model, regulating the necessary parameters of the trees.

Table 3: Tree-Specific Parameters

Parameter	Description
min_samples_split	Specifies the minimum number of samples required in a node.
min_samples_leaf	The minimum samples required in a terminal node or leaf.
min_weight_fraction_leaf	It is defined as a fraction of the total number of observations.
max_depth	Sets the maximum depth of a tree to control overfitting.
max_leaf_nodes	Limits the maximum number of terminal nodes or leaves in a tree.
max_features	Defines the number of features considered while searching for the best split.

Table 4 represents the boosting parameters used to decide the sequential modeling and even the contribution of each tree in the gradient-boosting textual algorithm. These parameters have a strong influence on the model complexity and capacity to learn the data while avoiding overfitting.

Table 4: Boosting Parameters

Parameter	Description
learning_rate	Governs the impact of each tree on the outcome.
n_estimators	Specifies the number of sequential trees to be modeled.
subsample	Represents the fraction of observations to be selected for each tree.

Table 5 below indicates the miscellaneous parameters used in tuning the Gradient Boosting algorithm. These parameters offer extra regulation concerning the training process and will alter the efficiency of the model training in some way. Knowledge of these parameters is essential for the best operations of the GB algorithm.

Table 5: Miscellaneous Parameters

Parameter	Description
loss	Refers to the loss function to be minimized in each split.
Init	Affects the initialization of the output.
random_state	The random number seed ensures reproducibility.
verbose	Dictates the type of output generated during the model fitting process.
warm_start	Allows additional trees to be fitted on the previous model fit.
presort	Determines whether to presort data for faster splits.

These parameters are very important in enhancing the efficiency of the GB algorithm, and hence, it is necessary to understand them. The next sections of the paper will explain how this DTO is incorporated with ML models and how it impacts the system's general performance.

4 Experimental Results

In this research study, we initially used the binary Dipper Throated Optimization (bDTO) for the feature selection. After this, we fine-tuned the cybersecurity threat detection using Dipper Throated Optimization (DTO) with Gradient Boosting. It was then compared to other metaheuristic optimization algorithms to assess the efficiency of the said approach. The results are shown in tables and figures, which makes it possible to understand the specifics of the interaction between feature selection methods and machine learning models.

4.1 Feature Selection Results

This paper focuses on feature selection in machine learning models, stating that feature selection is one of the most critical components that determine the performance of a machine learning model. The following table presents feature selection measures for different optimization algorithms. These are the average error, the average selected size, and the fitness indicators; the latter is indispensable for evaluating the efficacy of the algorithms in feature selection. Thus, having assessed the above-shown parameters, it is possible to identify comparative advantages and disadvantages of each of the algorithms and consequently choose the method most efficient for cybersecurity threat detection.

The average errors show how effectively each algorithm reduces wrong feature sets, and the average size of the selected set reveals the number of chosen features, which influences the model's complexity and readability. Average and worst fitness indicators represent the overall quality and stability of the feature selection process. Therefore, this Table serves as a conclusion for the initial assessment of an FCA.

Table 6: Results of the Feature Selection on the tested dataset

	bDTO	**bSCA**	**bMFO**	**bPSO**	**bWAO**	**bGWO**	**bFA**
Average error	0.64793	0.66513	0.67873	0.69893	0.69873	0.68523	0.69733
Average Select size	0.60073	0.80073	0.74313	0.80073	0.96413	0.72353	0.83523
Average Fitness	0.71113	0.72733	0.73873	0.72573	0.73353	0.73343	0.77763
Best Fitness	0.61293	0.64763	0.64203	0.70603	0.69763	0.71123	0.69633
Worst Fitness	0.71143	0.71453	0.75713	0.77373	0.77373	0.78743	0.79393
Standard deviation Fitness	0.53343	0.53813	0.53973	0.53753	0.53973	0.53873	0.57433

Figure 5 shows the performance of different feature selection techniques, comparing their accuracy levels. Such a chart facilitates comparing the results to identify which algorithms are more efficient in feature selection and provide the highest accuracy of the relevant features. In this way, the figure helps recognize which algorithm can be regarded as the most efficient in practice and offers clear references for further improvement and creation of another model.

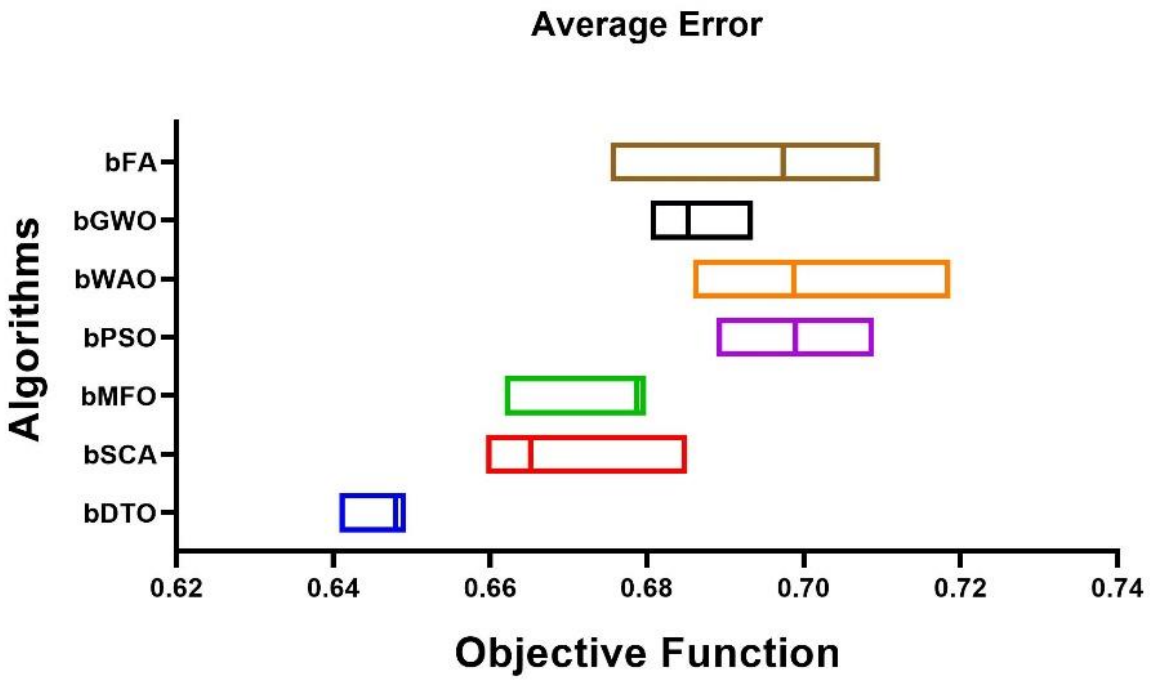

Figure 5: Accuracy investigation of various Feature Selection methods.

Further, compared with several other algorithms, including STDC, MoC, and Deep MRI, we plot the residual values and the heatmap for the bDTO algorithm in Figure 6. This visualization aids in visualizing the residual distribution and comparing each algorithm's efficacy to bDTO in feature selection. The residual values show how the predicted and the actual values differ, and the heatmap displays the variation across the numerous algorithms. Hence, from the figure, the researchers can capture the stability and reliability of each feature selection method; bDTO obtained minimum residual compared with other techniques proposed for lung cancer diagnosis.

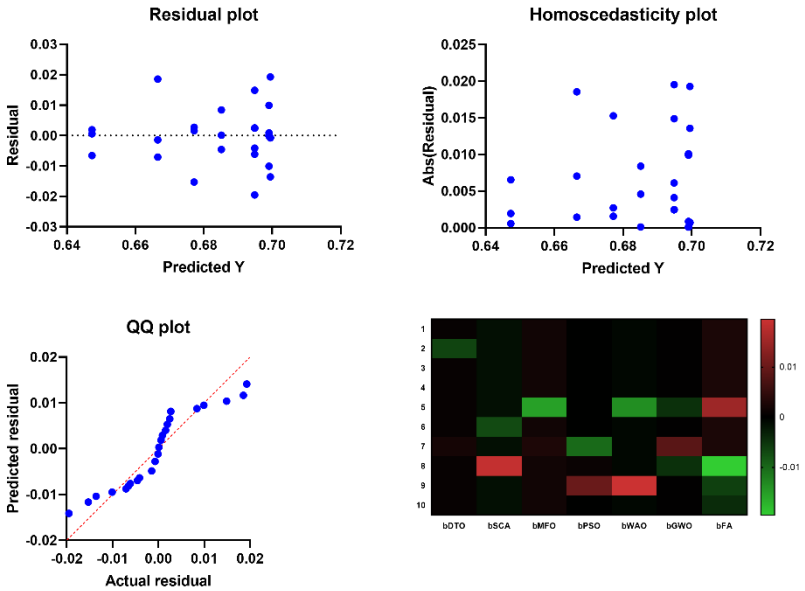

Figure 6: Residual Values and Heatmap for the proposed bDTO algorithm and other algorithms

Table 7 shows the ANOVA statistical values using the bDTO feature selection and the compared algorithms. In this analysis, the Sum of Squares (SS), Degrees of Freedom (DF), Mean Squares (MS), F-statistic, and p-value are necessary to comprehend the differences and variations of the several feature selection methods used.

The ANOVA results also show how different the values are from other algorithms, which is an essential method of comprehensive comparison of the bDTO algorithm's performance. The F-statistic is high and the P-value low, which means that bDTO has effectively increased the results, and therefore, the improvements registered cannot be attributed to external and random factors.

Table 7: ANOVA statistical results based on the bDTO and compared algorithms

ANOVA table	**SS**	**DF**	**MS**	**F (DFn, DFd)**	**P value**
Treatment (between columns)	0.02228	6	0.003714	F (6, 63) = 102.2	P<0.0001
Residual (within columns)	0.002289	63	0.00003633		
Total	0.02457	69			

Table 8 below depicts the Wilcoxon Signed Rank Test of the selected bDTO features and the compared algorithms. It presents the sum of the signed ranks, the positive ranks, the hostile ranks and the two-tailed p-value, which offers insight into how the paired samples tested differ in terms of the non-parametric Mann-Whitney U test.

Delving deep into the results, the p-value below 0.01 infuses more credence as to the substantial deviation in performance courtesy of the bDTO algorithm. The Wilcoxon test, being a nonparametric test, also indicates that the improvements with bDTO are indeed statistically significant and are unaffected by outliers or non-gaussian distributions.

Table 8: Wilcoxon Signed Rank Test statistical results for the bDTO Feature Selection and compared algorithms

Wilcoxon Signed Rank Test	**bDTO**	**bSCA**	**bMFO**	**bPSO**	**bWAO**	**bGWO**	**bFA**
The sum of signed ranks (W)	55	55	55	55	55	55	55
The sum of positive ranks	55	55	55	55	55	55	55
The sum of hostile ranks	0	0	0	0	0	0	0
P value (two-tailed)	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Is it exact or an estimate?	Exact	Exact	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**	**	**
Significant (alpha=0.05)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?							
Discrepancy	0.6479	0.6651	0.6787	0.6989	0.6987	0.6852	0.6973

4.2 Basic Machine Learning Models Results

Table 9 illustrates the basic machine learning models' performance metrics on the dataset. Different measuring scales, such as MSE, RMSE, MAE, and similar, are therefore valued to assess the accuracy of all the models.

Gradient Boosting, CatBoost and XGBoost models are highlighted as prospective since they can serve as a reference for comparison. The table's versatility enables the competition of each model, with specific values mentioning the areas of strength and areas where the given model can be improved.

Table 9: Results of the basic machine learning models on the tested dataset

Models	**MSE**	**RMSE**	**MAE**	**MBE**	**r**	**R2**	**RRM-SE**	**NSE**	**WI**	**Fitted Time**
GradientBoosting Regressor	0.0737	0.2715	0.2281	-0.0275	-0.1681	0.0283	58.3855	-0.0152	0.4962	0.0030
Cat Boost	0.0771	0.2778	0.2326	-0.0314	-0.1447	0.0209	59.7337	-0.0626	0.4862	11.6396
SVR	0.0776	0.2785	0.2332	-0.0124	-0.1315	0.0173	59.8906	-0.0682	0.4849	0.9253
XGBoost	0.0783	0.2797	0.2379	-0.0279	-0.1700	0.0289	60.1597	-0.0778	0.4744	11.6586
Linear Regression	0.0796	0.2821	0.2369	-0.0268	-0.0941	0.0088	60.6754	-0.0964	0.4768	0.0147
MLP Regressor	0.0819	0.2862	0.2362	-0.0257	-0.0386	0.0015	61.5577	-0.1285	0.4782	5.7622
Random Forest Regressor	0.0837	0.2893	0.2483	-0.0388	-0.1041	0.0108	62.2088	-0.1525	0.4516	1.3258
Decision Tree Regressor	0.0858	0.2929	0.2435	-0.0278	-0.1101	0.0121	62.9836	-0.1814	0.4622	0.1929
K Neighbors Regressor	0.0882	0.2970	0.2498	-0.0153	-0.1694	0.0287	63.8632	-0.2146	0.4483	0.0373
Extra Trees Regressor	0.0896	0.2993	0.2485	-0.0294	-0.1870	0.0350	64.3618	-0.2336	0.4511	0.0030

4.3 Optimization Results

Figure 7 presents the visual analysis of the accuracy related to research on DTO-Gradient Boosting and relatively more related algorithms. The given chart provides a clear vision of accuracy distribution, and once again, DTO-Gradient Boosting is in the leading position among other models. This kind of visualization helps understand the DTO-Gradient Boosting model's relative effectiveness and comprehend the optimization process's effects on the model's accuracy.

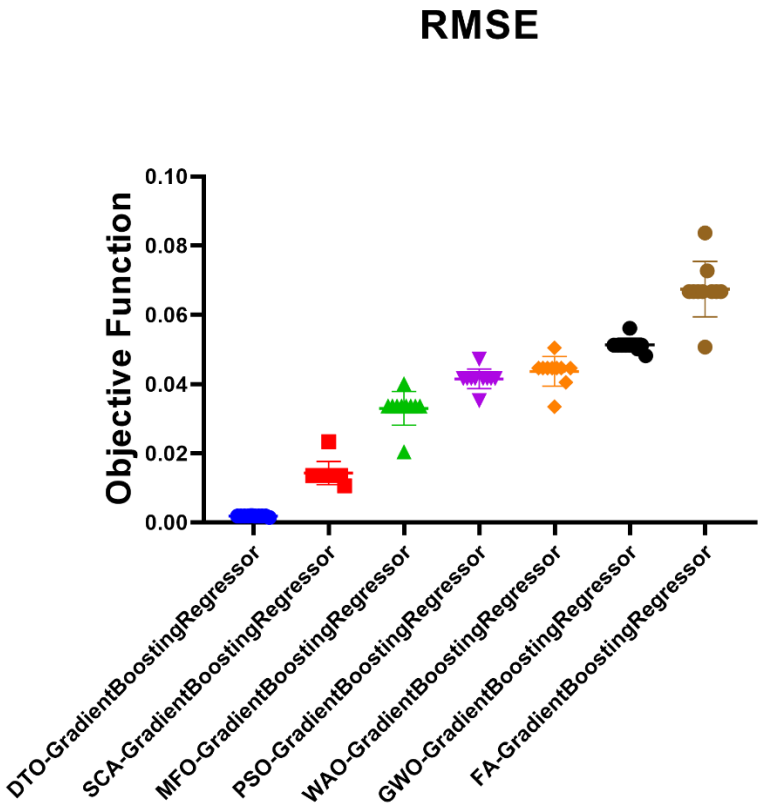

Figure 7: RMSE of DTO-GradientBoostingRegressor and compared algorithms

Figure 8 shows the histogram analysis of the accuracy of the implementation of DTO-Gradient Boosting and compared methods. This histogram aims to present the frequency of accuracy values, thus demonstrating how DTO has the potential to focus on highly accurate points. The result again confirms that DTO contributes positively to the optimization aspect of the gradient-boosting model. Based on the above figure, it is evident that the DTO-Gradient Boosting model is reliable and always has high accuracy, which makes it appealing for cybersecurity threat detection.

Figure 8: The Histogram of accuracy for DTO-GradientBoostingRegressor and other algorithms

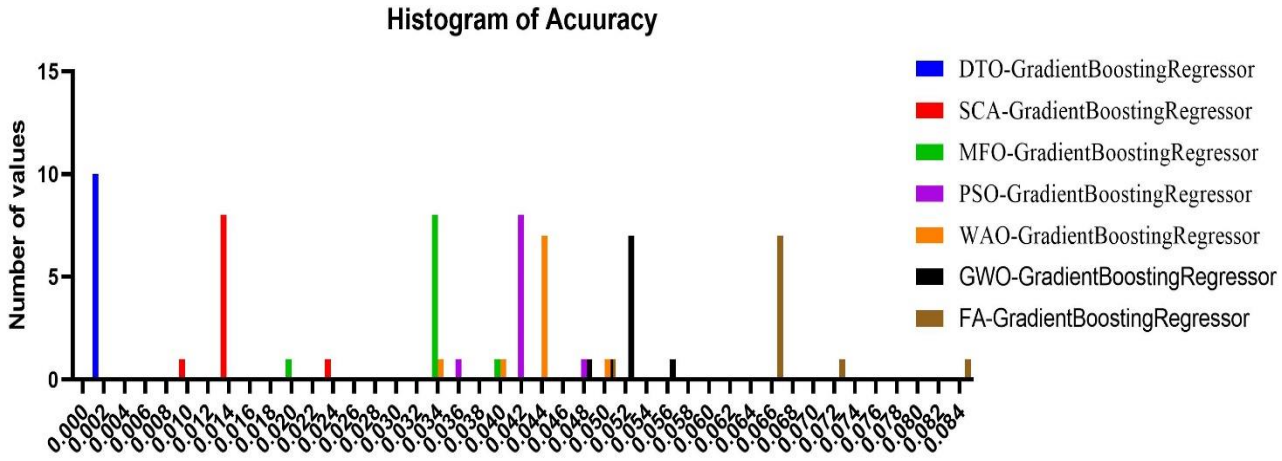

Table 10 encapsulates the performance and evaluation metrics summary for the DTO-Gradient Boosting model and the related comparison strategies. Statistical values like minimum, maximum, and Average give an indication of the variability and stability of the overall performance and models' performance. Using statistical measures that signify the DTO-Gradient Boosting model's credibility and efficiency is also helpful in establishing its degree of performance.

Table 10: Overview of the performance metrics of DTO-GradientBoostingRegressor models and compared algorithms on the tested dataset

	DTO	**SCA**	**MFO**	**PSO**	**WAO**	**GWO**	**FA**
# of values	10	10	10	10	10	10	10
Min.	0.001446	0.01058	0.02036	0.03516	0.03346	0.04812	0.05067
Median	0.001846	0.01358	0.03365	0.04157	0.04457	0.05124	0.06671
Max.	0.001995	0.02336	0.04004	0.04716	0.05046	0.05612	0.08367
Mean	0.00182	0.01426	0.03296	0.04149	0.04364	0.0513	0.0674
Std. Deviation	0.000139	0.003334	0.004859	0.002834	0.004299	0.001966	0.00799

The statistical values of the ANOVA test concerning the chosen DTO-Gradient Boosting model and the comparison of the algorithms are presented in Table 11. This F-statistic is relatively high ($p < 0.0001$) and supports the considerable difference in the performance of the two groups. This table gives a response of another type with statistical significance based on comparing the performance of the DTO Gradient Boosting and other algorithms. The significance of the findings is provided by the ANOVA results, which make the credibility of the change observed evident.

Table 11: ANOVA statistical results for the DTO-Gradient Boosting Regressor and other algorithms

	SS	**DF**	**MS**	**F (DFn, DFd)**	**P value**
Treatment (between columns)	0.02959	6	0.004931	F (6, 63) = 267.4	P<0.0001
Residual (within columns)	0.001162	63	0.0000184		
Total	0.03075	69			

Wilcoxon Signed Rank Test was conducted about DTO-Gradient Boosting (RMSE) and other algorithms, as shown in Table 12. The p-values for the tests are quite low at 0. 002 which shows that there are differences in the performances and confirms the findings from earlier tests whereby DTO-Gradient Boosting has the best performance. Wilcoxon test results also make up supporting evidence of the DTO-Gradient Boosting model for maximizing machine learning models for cybersecurity threat detection.

Table 12: Wilcoxon Test results for the DTO-Gradient Boosting Regressor and other algorithms

Wilcoxon Signed Rank Test	**DTO**	**SCA**	**MFO**	**PSO**	**WAO**	**GWO**	**FA**
Sum of signed ranks (W)	55	55	55	55	55	55	55
The sum of positive ranks	55	55	55	55	55	55	55
Sum of negative ranks	0	0	0	0	0	0	0
P value (two-tailed)	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Is it exact or an estimate?	Exact	Exact	Exact	Exact	Exact	Exact	Exact
P value summary	**	**	**	**	**	**	**
Significant (alpha=0.05)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
How big is the discrepancy?							
Discrepancy	0.001 846	0.013 58	0.03365	0.04157	0.04457	0.05124	0.06671

Lastly, Figure 9 demonstrates the residual values and the heatmap of the proposed algorithm alongside other algorithms in conjunction with gradient boost optimization. The heatmap offers a visually appealing way to view the residual distribution, thus stressing the versatility and credibility of DTO as an enhancing tool of the gradient-boosting model. The figure demonstrates how DTO helps reduce the amount of residuals and, as a result, can be beneficial for boosting the performance of cybersecurity threat detection models.

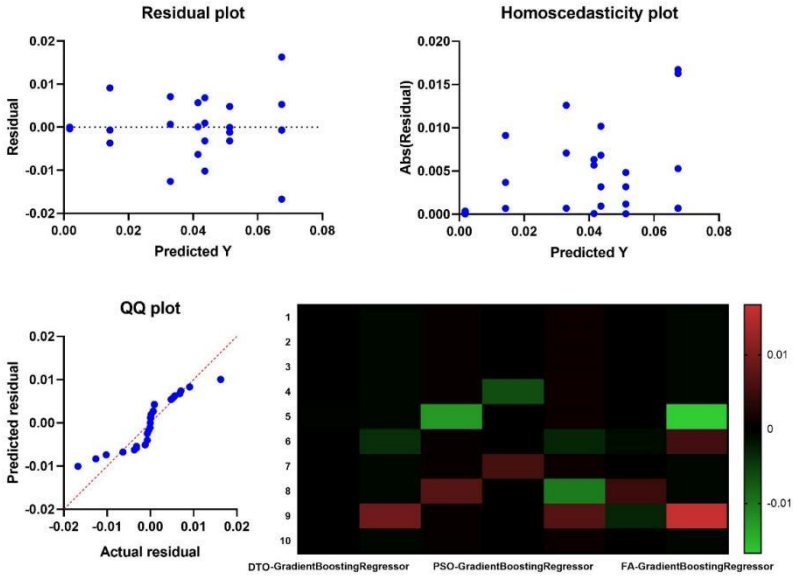

Figure 9: Residual Values and Heatmap for the proposed DTO algorithm and other algorithms

When comparing the DTO by way of the range and P values related to the various performance indicators and statistical tests, the efficacy and resilience of this work can be seen to be very high. Thus, it contributes to it being considered a valuable resource that might enhance the accuracy of the models used in machine

learning and help in threat recognition more efficiently. The subsequent sections of this paper will provide a preview of these results and offer directions for future work that might be helpful to scholars working in cyber defense and metaheuristic algorithms.

3. 5. Conclusions

In summary, this study researches the Binary Dipper Throated Optimization algorithm as an intelligent Integration component to find characteristics in optimizing the Gradient Boosting model for cybersecurity threat detection. Overall, bDTO operates in two modes. It turns out to be very fruitful when dealing with large datasets as this binary nature considerably increases the precision levels of features chosen for analysis. Based on a comparison with various performance metrics, Binary Dipper Throated Optimization consistently outperforms other algorithmic solutions. Besides the above, this study emphasizes how Dipper Throated Optimization techniques can be widely applicable in enhancing Gradient Boosting algorithms. In as significant a contribution to the optimization of machine learning models in cybersecurity, there is showcased here one potential capacity that harnesses from using the strengths, particularly its latent balance that comes about due to how it explores and exploits patterns within any given search space utilizing DTO. For future investigations, researchers could take this concept further and study the intricate capabilities of Dipper Throated Optimization because it focuses on binary optimization. Also, further studies could concentrate on even more complicated features of binary Dipper Throated Optimization and investigate how this algorithm can be adapted for different datasets related to cybersecurity. Furthermore, further studies of Dipper Throated Optimization methodologies would probably reveal new and novel ways in which a diverse array of machine learning algorithms might be optimized other than by Gradient Boosting. This research paves the foundation for future optimization methods and cybersecurity development with a platform that facilitates further innovations and practical use.In future work, adapting the DTO-Gradient Boosting framework for real-time deployment on resource-constrained IoT edge devices will be essential. This includes exploring lightweight variants of DTO or simplified optimization schemes that reduce computational demands. To improve adaptability, future research should also incorporate online learning techniques or adaptive mechanisms that can handle concept drift and evolving cyber threat patterns. Additionally, evaluating the system's robustness against adversarial attacks and deploying it in live IoT environments through pilot studies will be critical steps toward real-world implementation. These directions will help ensure the model's scalability, resilience, and operational feasibility.

Author Contributions: All authors have contributed equally

Funding: "This research received no external funding"

Informed Consent Statement: Not applicable

Availability of Data and Materials: The data that support the findings of this research are openly available at https://www.kaggle.com/datasets/teamincribo/cyber-security-attacks/

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. N. Vandezande, "Cybersecurity in the EU: how the nis2-directive stacks up against its predecessor," Computer Law & Security Review, vol. 52, p. 105890, 2024, doi: 1.1016/j.clsr.2023.105890.
2. Z. Turgut and A. G. Kakisim, "An explainable hybrid deep learning architecture for WIFI-based indoor localization in internet of things environment," Future Generation Computer Systems, vol. 151, pp. 196–213, 2024, doi: 1.1016/j.future.2023.1.003.
3. H. C. Leligou, A. Lakka, P. A. Karkazis, J. P. Costa, E. M. Tordera et al., "Cybersecurity in supply chain systems: the farm-to-fork use case," Electronics, vol. 13, no. 1, Art. no. 1, 2024, doi: 1.3390/electronics13010215.
4. N. Samee, E.-S. El-Kenawy, G. Atteia, M. Jamjoom, A. Ibrahim et al., "Metaheuristic optimization through deep learning classification of covid-19 in chest x-ray images," Computers, Materials & Continua, vol. 73, no. 2, pp. 4193–4210, 2022, doi: 1.32604/cmc.2022.031147.
5. A. M. Zaki, S. K. Towfek, W. Gee, W. Zhang and M. A. Soliman, "Advancing parking space surveillance using a neural network approach with feature extraction and dipper throated optimization integration," Journal of Artificial Intelligence and Metaheuristics, vol. Volume 6, no. Issue 2, pp. 16–25, 2023, doi: 1.54216/JAIM.060202.

6. W. Zhao, L. Wang, Z. Zhang, H. Fan, J. Zhang et al., "Electric eel foraging optimization: a new bio-inspired optimizer for engineering applications," Expert Systems with Applications, vol. 238, p. 122200, 2024, doi: 1.1016/j.eswa.2023.122200.
7. J. Zhang, L. Chen, Y. Xie, P. Yang, Z. Li et al., "Climate change mitigation in energy-dependent regions—a carbon tax-based cross-system bi-layer model with equilibrium-optimization superposition effects," Resources, Conservation and Recycling, vol. 200, p. 107315, 2024, doi: 1.1016/j.resconrec.2023.107315.
8. B. Aslani and S. Mohebbi, "Learn to decompose multiobjective optimization models for large-scale networks," International Transactions in Operational Research, vol. 31, no. 2, pp. 949–978, 2024, doi: 1.1111/itor.13169.
9. M. M. Eid, E.-S. M. El-Kenawy, N. Khodadadi, S. Mirjalili, E. Khodadadi et al., "Meta-heuristic optimization of LSTM-based deep network for boosting the prediction of monkeypox cases," Mathematics, vol. 10, no. 20, Art. no. 20, 2022, doi: 1.3390/math10203845.
10. A. M. Zaki, N. Khodadadi, W. H. Lim and S. K. Towfek, "Predictive analytics and machine learning in direct marketing for anticipating bank term deposit subscriptions," American Journal of Business and Operations Research, vol. Volume 11, no. Issue 1, pp. 79–88, 2023, doi: 1.54216/AJBOR.110110.
11. V. H. S. Pham, N. T. Nguyen Dang and V. N. Nguyen, "Enhancing engineering optimization using hybrid sine cosine algorithm with roulette wheel selection and opposition-based learning," Scientific Reports, vol. 14, no. 1, Art. no. 1, 2024, doi: 1.1038/s41598-024-51343-w.
12. A. K. Jain and L. Gidwani, "Dynamic economic load dispatch in microgrid using hybrid moth-flame optimization algorithm," Electrical Engineering, 2024, doi: 1.1007/s00202-023-02183-w.
13. E.-S. M. El-Kenawy, N. Khodadadi, S. Mirjalili, T. Makarovskikh, M. Abotaleb et al., "Metaheuristic optimization for improving weed detection in wheat images captured by drones," Mathematics, vol. 10, no. 23, Art. no. 23, 2022, doi: 1.3390/math10234421.
14. D. Khafaga, A. Alhussan, E.-S. El-kenawy, A. Ibrahim, S. H et al., "Improved prediction of metamaterial antenna bandwidth using adaptive optimization of lstm," Computers, Materials & Continua, vol. 73, no. 1, pp. 865–881, 2022, doi: 1.32604/cmc.2022.028550.
15. M. A. Hassan, N. Bailek, K. Bouchouicha, A. Ibrahim, B. Jamil et al., "Evaluation of energy extraction of pv systems affected by environmental factors under real outdoor conditions," Theoretical and Applied Climatology, vol. 150, no. 1, pp. 715–729, 2022, doi: 1.1007/s00704-022-04166-6.
16. A. Djaafari, A. Ibrahim, N. Bailek, K. Bouchouicha, M. A. Hassan et al., "Hourly predictions of direct normal irradiation using an innovative hybrid lstm model for concentrating solar power projects in hyper-arid regions," Energy Reports, vol. 8, pp. 15548–15562, 2022, doi: 1.1016/j.egyr.2022.1.402.
17. S. S. Bagalkot, D. H. A and N. Naik, "Novel grey wolf optimizer based parameters selection for garch and arima models for stock price prediction," PeerJ Computer Science, vol. 10, p. e1735, 2024, doi: 1.7717/peerj-cs.1735.
18. M. Karthikeyan, D. Manimegalai and K. RajaGopal, "Firefly algorithm based wsn-iot security enhancement with machine learning for intrusion detection," Scientific Reports, vol. 14, no. 1, Art. no. 1, 2024, doi: 1.1038/s41598-023-50554-x.
19. M. Macedo, H. Siqueira, E. Figueiredo, C. Santana, R. C. Lira et al., "Overview on binary optimization using swarm-inspired algorithms," IEEE Access, vol. 9, pp. 149814–149858, 2021, doi: 1.1109/ACCESS.2021.3124710.
20. F. H. Rizk, S. Arkhstan, A. M. Zaki, M. A. Kandel and S. K. Towfek, "Integrated cnn and waterwheel plant algorithm for enhanced global traffic detection," Journal of Artificial Intelligence and Metaheuristics, vol. Volume 6, no. Issue 2, pp. 36–45, 2023, doi: 1.54216/JAIM.060204.
21. H. AlEisa, E.-S. El-kenawy, A. Alhussan, M. Saber, A. Abdelhamid et al., "Transfer learning for chest x-rays diagnosis using dipper throated algorithm," Computers, Materials & Continua, vol. 73, no. 2, pp. 2371–2387, 2022, doi: 1.32604/cmc.2022.030447.
22. A. Abdelmgeed, A. M. Zaki and M. A. Soliman, "An evaluation of ARIMA and persistence models in iot-driven smart homes," Journal of Artificial Intelligence and Metaheuristics, vol. Volume 6, no. Issue 2, pp. 08–15, 2023, doi: 1.54216/JAIM.060201.
23. E.-S. El-kenawy, A. Ibrahim, S. Mirjalili, Y.-D. Zhang, S. Elnazer et al., "Optimized ensemble algorithm for predicting metamaterial antenna parameters," Computers, Materials & Continua, vol. 71, no. 3, pp. 4989–5003, 2022, doi: 1.32604/cmc.2022.023884.
24. M. Arumugam, A. Thiyagarajan, L. Adhi and S. Alagar, "Crossover smell agent optimized multilayer perceptron for precise brain tumor classification on mri images," Expert Systems with Applications, vol. 238, p. 121453, 2024, doi: 1.1016/j.eswa.2023.121453.
25. M. Alviano, F. Bartoli, M. Botta, R. Esposito, L. Giordano et al., "A preferential interpretation of multilayer perceptrons in a conditional logic with typicality," International Journal of Approximate Reasoning, vol. 164, p. 109065, 2024, doi: 1.1016/j.ijar.2023.109065.
26. M. Al-Omari, M. Rawashdeh, F. Qutaishat, M. Alshira'H and N. Ababneh, "An intelligent tree-based intrusion detection model for cyber security," Journal of Network and Systems Management, vol. 29, no. 2, p. 20, 2021, doi: 1.1007/s10922-021-09591-y.
27. L. Haghnegahdar and Y. Wang, "A whale optimization algorithm-trained artificial neural network for smart grid cyber intrusion detection," Neural Computing and Applications, vol. 32, no. 13, pp. 9427–9441, 2020, doi: 1.1007/s00521-019-04453-w.
28. I. H. Sarker, M. H. Furhad and R. Nowrozy, "AI-driven cybersecurity: an overview, security intelligence modeling and research directions," SN Computer Science, vol. 2, no. 3, p. 173, 2021, doi: 1.1007/s42979-021-00557-0.

29. N. N. Abbas, T. Ahmed, S. H. U. Shah, M. Omar and H. W. Park, "Investigating the applications of artificial intelligence in cyber security," Scientometrics, vol. 121, no. 2, pp. 1189–1211, 2019, doi: 1.1007/s11192-019-03222-9.
30. S. Zhang, X. Xie and Y. Xu, "A brute-force black-box method to attack machine learning-based systems in cybersecurity," IEEE Access, vol. 8, pp. 128250–128263, 2020, doi: 1.1109/ACCESS.202.3008433.
31. I. H. Sarker, "CyberLearning: effectiveness analysis of machine learning security modeling to detect cyber-anomalies and multi-attacks," Internet of Things, vol. 14, p. 100393, 2021, doi: 1.1016/j.iot.2021.100393.
32. D. Wang, X. Wang, Y. Zhang and L. Jin, "Detection of power grid disturbances and cyber-attacks based on machine learning," Journal of Information Security and Applications, vol. 46, pp. 42–52, 2019, doi: 1.1016/j.jisa.2019.02.008.
33. Y. Jiang and Y. Atif, "A selective ensemble model for cognitive cybersecurity analysis," Journal of Network and Computer Applications, vol. 193, p. 103210, 2021, doi: 1.1016/j.jnca.2021.103210.
34. "Cyber security attacks." Accessed: Jan. 12, 2024. [Online]. Available: https://www.kaggle.com/datasets/teamincribo/cyber-security-attacks
35. Z. Liu, P. Jiang, K. W. De Bock, J. Wang, L. Zhang et al., "Extreme gradient boosting trees with efficient bayesian optimization for profit-driven customer churn prediction," Technological Forecasting and Social Change, vol. 198, p. 122945, 2024, doi: 1.1016/j.techfore.2023.122945.
36. L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: catboost and xgboost approaches," Expert Systems with Applications, vol. 241, p. 122686, 2024, doi: 1.1016/j.eswa.2023.122686.
37. X. Ma, J. Yang, R. Zhang, W. Yu, J. Ren et al., "XGBoost-based analysis of the relationship between urban 2d/3d morphology and seasonal gradient land surface temperature," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, pp. 1–17, 2023, doi: 1.1109/JSTARS.2023.3348476.
38. S. Hoffman and R. Jasiński, "The use of multilayer perceptrons to model pm2.5 concentrations at air monitoring stations in poland," Atmosphere, vol. 14, no. 1, Art. no. 1, 2023, doi: 1.3390/atmos14010096.
39. H. Rasyid, N. Hariani Soekamto, Seniwati, S. Firdausiah and Firdaus, "Modelling the anticancer activity of 4-alkoxy cinnamic analogues using 3d-descriptors and genetic algorithm-multiple linear regression (ga-mlr) method," Journal of King Saud University - Science, vol. 35, no. 3, p. 102514, 2023, doi: 1.1016/j.jksus.2022.102514.
40. Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang et al., "An improved random forest based on the classification accuracy and correlation measurement of decision trees," Expert Systems with Applications, vol. 237, p. 121549, 2024, doi: 1.1016/j.eswa.2023.121549.
41. D. S. Khafaga, A. Ibrahim, E.-S. M. El-Kenawy, A. A. Abdelhamid, F. K. Karim et al., "An al-biruni earth radius optimization-based deep convolutional neural network for classifying monkeypox disease," Diagnostics, vol. 12, no. 11, Art. no. 11, 2022, doi: 1.3390/diagnostics12112892.
42. E.-S. M. El-kenawy, B. Zerouali, N. Bailek, K. Bouchouich, M. A. Hassan et al., "Improved weighted ensemble learning for predicting the daily reference evapotranspiration under the semi-arid climate conditions," Environmental Science and Pollution Research, vol. 29, no. 54, pp. 81279–81299, 2022, doi: 1.1007/s11356-022-21410-8.
43. M. Karbasi, M. Ali, S. M. Bateni, C. Jun, M. Jamei et al., "Boruta extra tree-bidirectional long short-term memory model development for pan evaporation forecasting: investigation of arid climate condition," Alexandria Engineering Journal, vol. 86, pp. 425–442, 2024, doi: 1.1016/j.aej.2023.11.061.
44. A. A. Abdelhamid, E.-S. M. El-Kenawy, A. Ibrahim, M. M. Eid, D. S. Khafaga et al., "Innovative feature selection method based on hybrid sine cosine and dipper throated optimization algorithms," IEEE Access, vol. 11, pp. 79750–79776, 2023, doi: 1.1109/ACCESS.2023.3298955.
45. M. Y. Shams, E.-S. M. El-kenawy, A. Ibrahim and A. M. Elshewey, "A hybrid dipper throated optimization algorithm and particle swarm optimization (DTPSO) model for hepatocellular carcinoma (HCC) prediction," Biomedical Signal Processing and Control, vol. 85, p. 104908, 2023, doi: 1.1016/j.bspc.2023.104908.