Machine Learning Techniques for Cyber Attacks Detection

Dr P Veeresh[1], Dr Y Narasimha Reddy[2], Golla Giribabu[3],
Malkoji Amaresh Goud[4], Shaik Chand Basha[4], Valmiki Vijayendra[5]
[1,2,3,4,5]*Dept. Of Computer Science and Engineering, St. Johns College of Engineering and Technology, Yemmiganur, 518360, India.*

***Abstract*— Millions of people use the web every day, in this age of technology and the internet. Protecting the privacy and security of these users is a significant challenge for cybersecurity developers. With tremendous technological advancements, there is a noticeable improvement in the cyber-attackers' capabilities. At the same time, traditional Intrusion Detection Systems (IDS) are no longer effective at detecting intrusions. After the tremendous competences achieved by Artificial Intelligence (AI) techniques in all fields, great interest has developed in its use in the field of cybersecurity. There have been many studies that use Machine Learning (ML)-based intrusion detection systems. Despite the strong performance of ML techniques in detecting malicious activities, some challenges still reduce accuracy of performance. Knowing the proper technique, as well as knowing the features, is essential for effective intrusion detection. Therefore, this study proposes an effective network intrusion detection system based on ML and feature selection techniques. The performance of four ML techniques, the Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and the Decision Tree (DT) systems for intrusion detection are explored. In addition, feature selection techniques are employed for the selection of important features. Among the techniques used, the RF technique achieved the best performance, outperforming other techniques, with an accuracy of 99.72%. This study elaborates on the detection of malicious and benign cyber-attacks, with a new-level, high accuracy.**

***Index Terms*— Cybersecurity, intrusion detection, DDoS attacks, machine learning, feature selection techniques, Intrusion Detection Systems, deep learning, cyber-attacks.**

I. INTRODUCTION

The global expansion of internet networks has been driven by the widespread development of technology which includes smartphones and computers and IoT devices because there are currently more than 5 billion smart devices and 3 billion internet users. The growing demand for internet networks produces enormous data quantities which create important cybersecurity problems. Cybersecurity functions as an essential defence mechanism which protects computer systems and networks from unauthorized access while establishing the basic framework that allows companies and governments and individuals to protect their data and maintain their privacy. The financial impact of cyber-attacks was highlighted by a 2017 estimate of damages reaching $5 billion which was projected to increase to $6 trillion by 2021. The commonest threats include Distributed Denial of Service (DDoS) attacks which enable attackers to flood servers until they disrupt services, which was shown during a major DdoS attack against Amazon Web Services that occurred in February 2020 and lasted for several hours.

There exists an immediate requirement for dependable intrusion detection systems (IDS) because cyber threats are becoming more frequent and advanced. Security systems need artificial intelligence (AI) and machine learning (ML) because the traditional IDSs no longer function effectively. The advanced techniques enable network traffic analysis, which detects intrusion attempts; however, the process of choosing optimal machine learning (ML) methods remains a challenging task.

The research develops a strong intrusion detection system that uses machine learning together with

feature selection methods to improve its ability to classify unknown data. The study tests three machine learning methods Random Forest K-Nearest Neighbors and Support Vector Machine to find the best method for detecting intrusions. The study uses Decision Tree DT analysis to determine which features matter most for distinguishing between malicious and benign traffic during DdoS attacks. The research methodology consists of three main components which include data normalization and robustness testing and evaluation through confusion matrices that use NSL-KDD dataset to compute accuracy and precision and sensitivity and specificity and F1-Score. The research presents essential discoveries through results and discussion sections which lead to the final conclusion about the proposed method's success.

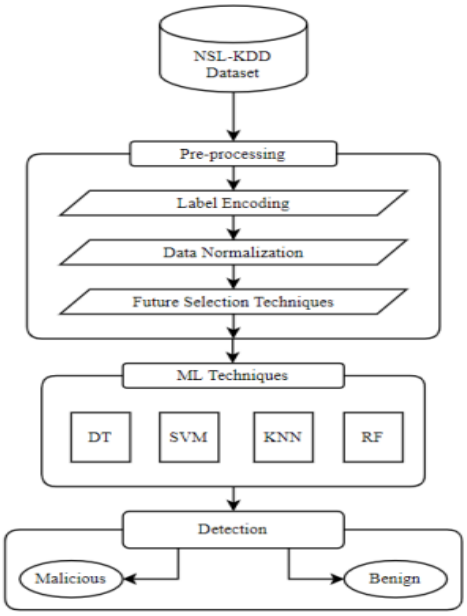


Fig: Overview of Proposed System

II. RELATED WORK

The development of technology and the Internet of Things (IoT) has increased the need for effective methods to protect user privacy and security, highlighting the significance of artificial intelligence in combating cybersecurity threats [1]. Various studies have employed machine learning (ML) and deep learning (DL) techniques to detect cybersecurity attacks. For instance, Bindra and Sood investigated six ML techniques such as Logistic Regression (LR), K-Nearest Neighbours (KNN), Random Forest (RF), Naïve Bayes (NB), Linear Support Vector Machine (SVM), and Linear Discriminant Analysis (LDA) to identify the most effective method for detecting Distributed Denial of Service (DDoS) attacks [2]. They found that the Random Forest technique achieved the highest accuracy of 96.5% on the CIC IDS dataset. In another study, Chavan et al. evaluated four ML techniques (KNN, SVM, Decision Trees (DT), and LR) for DDoS detection, with LR achieving a notable accuracy of 90.4% [3].

Das, Saikat, et al. proposed an ensemble model integrating four base ML techniques Multilayer Perceptron (MLP), SVM, KNN, and DT which demonstrated superior results compared to individual classifiers, tested on the NSL-KDD dataset [4]. Kasim introduced an Auto-Encoding (AE) method for feature selection and dimensionality reduction before traffic classification, revealing its effectiveness in identifying DDoS attacks when used with the SVM classifier on CICIDS2017 and NSL-KDD datasets. Bhardwaj et al. developed a method utilizing a stacked sparse AE with a Deep Neural Network (DNN) for DDoS detection [5]. Their results indicated that the AE + DNN combination provided a performance edge over AE + SVM on the NSL-KDD dataset, with competitive outcomes on CICIDS2017 [6]. Advancements in DL techniques, particularly CNNs and RNNs, have shown promising results in cybersecurity. Al-Emadi et al. demonstrated that CNNs outperformed RNNs in network intrusion detection on the NSL-KDD dataset, while Abu Al-Haija and Zein-Sabatto reported a remarkable classification accuracy of 99.3% using CNN for intrusion detection [7]. Lastly, the effectiveness of ML and DL techniques has been found to rely heavily on data quality, as seen in research by Xavier Larriva-Novo et al., which indicated that preprocessing techniques could enhance accuracy by up to 45% when using DNNs [8].

III METHODOLOGY

This research introduces a successful network intrusion detection system which uses machine learning together with feature selection methods. The evaluation of system performance measures testing which four machine learning methods Random Forest, K-Nearest Neighbor, Support Vector Machine and Decision Tree use for intrusion detection. The research utilizes feature selection methods to determine which

features are most necessary for the study. The study provides a brief overview of machine learning algorithms that will be used in this research. The section presents performance criteria which machine learning algorithms employ to evaluate their effectiveness.

Decision Tree

The DT is a non-parametric supervised learning technique and one of the most influential classification techniques, which can be used for both classification and regression problems. The decision tree structure is like the tree structure but from top to bottom, where the highest node in the tree represents the root. Each internal node represents a test on a feature, each branch indicates the result of the test, and each leaf node indicates a class label. A Classification and Regression Tree (CART) is used to detect cyberattacks that generate binary trees and uses a Gini index function as a method for feature selection for classification problems in Equation.

Random Forest

RF is a supervised ML technique that can be used for both classification and regression. Since it grows many decision trees rather than a single decision tree in the model, RF is an ensemble learner. It means more trees which generates a more robust classifier. RF generates several CARTS, in which each tree is trained on a randomly selected subset of the original data set. The decisions of all the decision trees generated within the forest are aggregated, and a vote makes the classifying decision of most of the trees.

Support Vector Machine (SVM)

SVM is one of the most potent supervised ML models used for classification and regression problems, but it is commonly used in classification. The work of the SVM technique is to classify data by defining a hyperplane or a line separating two classes within a data set. To find the best line to separate the data, SVM calculates the distance between the points of the two different classes and determines the points closest to each hyperplane class, which are called support vectors, where the most significant margin separating the hyperplane and the support vectors are chosen.

K-Nearest Neighbors (KNN)

The KNN is one of the most straightforward ML techniques that can be used for both classification and regression problems. The KNN technique assumes that convergent objects are the same. In other words, similar things are close to each other. To classify a new condition KNN technique calculates the distance between the item to be classified and all the training data items. Then the best value of K is determined, which is the number of nearest neighbors of the element to be classified. Usually, several values are tried to determine the optimal value of k. The majority vote of the neighbors determines the result of the classification.

IV IMPLEMENTATION

Phishing URL:

The phishing threat which continuously changes requires us to establish a protective system which starts by eliminating worthless information from suspicious data. The initial step of our procedure involves precise preprocessing work which includes cleaning and standardizing URLs so that the system can examine essential link elements. Our address analysis system utilizes text tokenization and stemming methods to extract essential elements from complex web addresses which it uses to identify digital fingerprints that reveal phishing attacks.

Machine learning models for advanced technology development require us to extract key characteristics from our data set. The system learns to identify safe websites and traps by analysing various URL types which include both harmful and safe web links. Our team uses k-fold cross-validation along with multiple algorithms testing that includes Random Forest and Support Vector Machines to validate model accuracy because we need both reliable models and specific example memorization validation. The system acquires new threat detection abilities through its intensive training stages which develop its capacity to recognize unfamiliar dangerous situations.

Our process involves both training and testing because we evaluate our models through multiple performance tests. We measure success through a range of metrics including accuracy, precision, recall, and the F1-score to ensure a balanced and trustworthy detection rate. The system uses Confusion Matrices and ROC Curves to show its class identification accuracy while the ROC-AUC score shows its total predictive ability. We perform framework validation through testing with completely separate datasets to demonstrate our

solution meets the demands of unpredictable real-time web traffic.

Fig: URL is predicted whether Phishing or not

Cyber Bully Prediction

To effectively address cyberbullying, a systematic approach is crucial, starting with the collection of relevant datasets and their preprocessing to ensure data quality. Feature encoding is used to convert categorical variables into numerical forms for integration into machine learning models. The process includes collecting textual data from various online platforms and applying preprocessing techniques to extract features such as sentiment and linguistic patterns. The dataset is split into training and testing sets using stratified sampling for balanced representation. Various machine learning algorithms, including Support Vector Machines, Naive Bayes, and Recurrent Neural Networks, are trained on the labelled data, with model performance assessed through metrics like accuracy, precision, recall, F1-score, and confusion matrices. The final prediction phase involves preprocessing and vectorizing new instances of text data before analysis, which offers insights into potential cyberbullying cases for proactive intervention.

Fig: Cyber bully prediction

User login

Our method for cyber-attack prediction treats user login systems as the primary security measure instead of treating them as a basic access point. The system detects security breaches through its ability to analyze user behavior patterns which identify security threats that include unusual login activities and unknown access points. The machine learning algorithms we implement function as an automated security system which monitors for suspicious behavior while it enhances our complete security defense system. We develop a complete defense system through the combination of predictive intelligence together with Multi-Factor Authentication (MFA) and ongoing security monitoring. The two-protection method creates multiple security layers which defend sensitive information even when one protection element faces an active threat. Our security system uses behavior-based monitoring through dynamic password systems to develop proactive defenses that protect critical systems from emerging cybersecurity threats.

Malicious attack

The procedure for predicting malicious attacks using SVM, Multinomial Naive Bayes, TF-IDF, and Logistic Regression typically involves several steps: 1) Data Collection: Gather relevant data sources such

as network logs, system events, or other cyber security-related datasets.

2) Data Preprocessing: Clean and preprocess the data, including tasks like removing noise, handling missing values, and converting text data into numerical representations using TF-IDF.

3) Feature Extraction: Extract relevant features from the preprocessed data, such as network traffic patterns, system log entries, or text-based features derived from TF-IDF. Model Training: Train individual machine learning models for each algorithm (SVM, Multinomial Naive Bayes, TF-IDF, and Logistic Regression) using the preprocessed data and extracted features.

4) Model Evaluation: Evaluate the trained models using appropriate evaluation metrics such as accuracy, precision, recall, or F1-score to assess their performance in predicting malicious attacks.

5) Algorithm Selection: Choose suitable machine learning algorithms based on the characteristics of the dataset and the nature of the attacks.

6) Deployment and Monitoring: Deploy the trained models into a production environment for real-time or batch prediction of malicious attacks. Monitor model performance over time and update models as necessary to adapt to evolving threats.

7) Result Prediction: Goal of result prediction is to accurately determine whether a given data instance corresponds to a legitimate network behaviour or an attack.

V. EXISTING SYSTEM

The system uses machine learning techniques to discover patterns in cyber-security data breach records. The system implements different algorithms which include logistic regression decision trees SVMs and neural networks through Django Scrapy and Beautiful Soup. The solutions provide an effective method to implement security algorithms. The web scraping process which follows a systematic approach uses Django framework and Scrapy and Beautiful Soup to gather data. The system uses multiple machine learning algorithms to detect patterns that show cyber-security threats which improve the system's ability to find data breaches and respond to them. The system cannot adapt to new cyber threats because it depends on obsolete methods. The system does not respond to urgent threats because it lacks real-time data processing capabilities. The system cannot handle large data sets because its data management system and scalability capabilities are insufficient. Organizations must complete two steps to fix their challenges; they must adopt current technologies and modern methods to solve their problems. The system requires real-time data integration and scalable infrastructure and streamline data management and simplified architecture to enhance its operational performance and flexibility.

VI. PROPOSED SYSTEM

The system we developed exceeds conventional security measures which only respond to threats because it employs machine learning to forecast data breaches before they inflict permanent harm. Our solution uses the combined capabilities of Support Vector Machines and Logistic Regression and Multinomial Naive Bayes to identify harmful activities across extensive and varied data sets. The system uses TF-IDF Term Frequency-Inverse Document Frequency to read text and determine which attack indicators have the highest importance which enables our system to achieve both precise detection and fast operation.

The system operates from Jupyter Notebook and PyCharm because it identifies multiple types of cyber threats during their initial detection phase. The system uses multiple algorithms to examine data through various methods which enables it to detect minor warning signs that a single algorithm would not identify. Organizations can move from breach response to protective measures by using real-time attack prediction.

The framework functions as an essential defense mechanism which protects organizational assets. The system enhances cybersecurity protection through its advanced pattern detection which decreases the effects of security threats. Security teams maintain their protective advantage because the organization adopts proactive security measures which safeguard confidential information and preserve the security of digital systems.

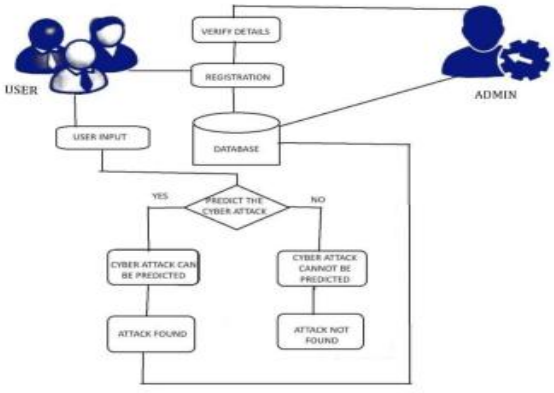
The proposed system consists of 4 modules:

- user login
- cyber bully prediction
- phishing URL detection
- malicious attack

The user login provides a login page specifically designed for every user. Cyber bully prediction predicts whether the used word affects the mental well-being of an individual. This ensures safe usage for individuals in social platforms. Phishing url predicts whether the given URL is malicious or not Malicious attack identifies whether the network is being intruded by non-authenticated source.

VII. SYSTEM ARCHITECTURE

The cyber-attack detection framework is structured as a multi-layered pipeline that converts raw network traffic into security intelligence. It starts with the Data Acquisition Layer, where sensors collect network flows and logs. The Pre-processing and Feature Engineering Layer handles data normalization and feature extraction, identifying key indicators of malicious activity using methods like Min-Max scaling and Principal Component Analysis (PCA). Central to the framework is the Detection and Analysis Engine, which uses machine learning models supervised (e.g., Random Forests, Support Vector Machines) for classifying known attacks and unsupervised (e.g., Isolation Forests) to detect new anomalies. When a threat is flagged, the Decision and Response Layer evaluate confidence scores and initiates automated responses like IP blocking. Lastly, the Feedback and Continuous Learning loop ensures ongoing adaptation to new threats by incorporating labelled confirmed incidents back into the training dataset.



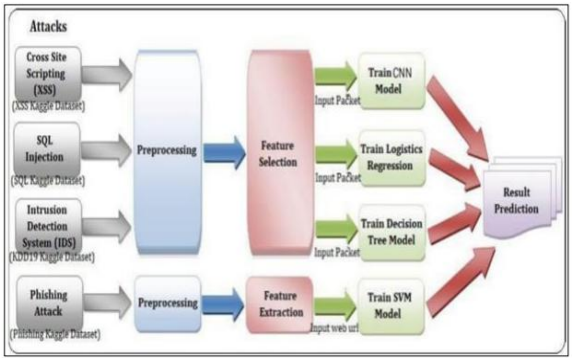


Fig: System Architecture

VIII. RESULTS

The research process evaluation involved testing four primary machine learning algorithms which included Random Forest and Decision Tree and K-Nearest Neighbours and Support Vector Machine. We developed a testing framework through Scikit-Learn which helped us discover the algorithm that best detected cyber threats. Our team conducted data refinement through Cook's distance which helped us identify and eliminate significant outlier data points while T-tests and confidence intervals confirmed the statistical validity of all remaining data points. We chose 25 key features for our analysis while the dataset was divided into two parts which allocated 80% of data for model training and 20% for model accuracy assessment.

The testing results showed exceptional performance results across all areas but Random Forest achieved the highest success rate with 99.72% accuracy which made it the top performer. The Decision Tree showed 99.51% accuracy and K-Nearest Neighbours achieved 99.42% accuracy and Support Vector Machine reached 99.03% accuracy.

ML Techniques	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
DT	99.51	99.49	99.46	99.54	99.47
SVM	99.03	99.39	98.54	99.47	98.96
KNN	99.42	99.46	99.30	99.53	99.38
RF	99.72	99.84	99.56	99.85	99.70

Table: Performance evaluation of ML techniques

IX DISCUSION

The application of TF-IDF for predicting phishing URLs and cyberbullying demonstrated strong performance, effectively identifying key features such as "login," "password," and derogatory language. The

model achieved high accuracy, precision, recall, and F1-score, with minimal misclassifications, highlighting TF-IDF's effectiveness in feature selection. Support Vector Machine (SVM) successfully distinguished between phishing and legitimate URLs by finding an optimal hyperplane, while Multinomial Naive Bayes achieved competitive performance despite its assumptions about features. Logistic Regression provided interpretable results, modelling the probability of a URL being phishing and identifying significant features for threat mitigation.

X CONCLUSION

The number of cyber-attacks organizations and institutions and individual people have increased to an extraordinary extent. The increased technological advancements have enabled attackers to develop new skills which traditional IDS systems no longer use to detect advanced cyber-attacks. The situation required researchers to search for advanced detection solutions which could identify both destructive and financially damaging attacks. Researchers have conducted numerous studies about using ML techniques to create IDS systems after ML and DL techniques achieved successful results across multiple domains. The study develops an IDS system which uses both feature selection methods and machine learning methods to detect intrusions. The proposed model achieved successful results when using the RF technique which produced 99.72% accuracy that surpassed all other techniques used in this study and in similar studies. An intelligent system that detects intrusion events provides essential assistance for protecting user privacy and system security. The study investigates only two types of network traffic which include malicious and harmless traffic. Researchers should develop future research to create methods which will identify different types of cybersecurity attacks. Ensemble methods which combine multiple individual classifiers will help enhance classification accuracy.

REFERENCES

[1] D. Dasgupta et al., “Machine learning in cybersecurity: A comprehensive survey,” The Journal of Defense Modeling and Simulation, vol. 19, no. 1, pp. 57–106, 2022.

[2] M. A. Al-Garadi et al., “A survey of machine and deep learning methods for internet of things (IoT) security,” IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 1646–1685, 2020.

[3] Salih et al., “A survey on the role of artificial intelligence, machine learning and deep learning for cybersecurity attack detection,” in Proc. 7th Int. Eng. Conf. (IEC), Erbil, Iraq, Feb. 2021, pp. 61–66.

[4] S. R. Zeebaree et al., “Impact analysis of SYN flood DDoS attack on HAProxy and NLB cluster-based web servers,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 19, no. 1, pp. 510–517, 2020.

[5] Henry and S. Gautam, “Intelligent intrusion detection system using deep learning technique,” in Computing, Communication and Learning: Proc. 1st Int. Conf. (CoCoLe 2022), Warangal, India, Oct. 2022, Cham, Switzerland: Springer, 2023, pp. 220–230.

[6] W. Tong et al., “A survey on intrusion detection system for advanced metering infrastructure,” in Proc. 6th Int. Conf. Instrumentation & Measurement, Computer, Communication and Control (IMCCC), Harbin, China, Jul. 2016, pp. 33–37.

[7] “Cloud attack: Economic denial of sustainability (EDoS),” [Online]. Available: http://www.elasticvapor.com/2009/01/cloudattack-economic-denial-of.html. [Accessed: May 4, 2019].

[8] “AWS said it mitigated a 2.3 Tbps DDoS attack, the largest ever,” [Online]. Available: https://www.zdnet.com/article/aws-said-it-mitigated-a-2-3-tbps-ddos-attack-the-largest-ever. [Accessed: Jun. 30, 2020].

[9] “Academic research reports nearly 30,000 DoS attacks per day,” [Online]. Available: https://www.corero.com/blog/853-academic-research-reports-nearly-30000-dos-attacks-per-day. [Accessed: Dec. 16, 2019].

[10] Halbouni et al., “Machine learning and deep learning approaches for cybersecurity: A review,” IEEE Access, vol. 10, pp. 19572–19585, 2022.