DEEP LEARNING APPROACHES FOR IDENTIFYING CYBER ATTACKS IN DATA SCIENCE

Abstract

Cyber-attacks are a growing hazard to cyber- physical systems, and firms that handle sensitive data must be able to identify and anticipate these attacks. Because the types of Cyber Attacks are becoming more diverse, anti-virus scanners alone are no longer adequate to offer protection. The growth of Internet of Things (IoT) devices has led to significant increases in cyber security vulnerabilities, and the market for software is growing as more software is used in more aspects of daily life. Hackers launch cyber-attacks in a variety of ways, including Phishing, Dos, R2L, Probing, Malware, and U2R. Large data sets can occasionally be compromised without the affected parties' knowledge, which can cause serious interruptions to business continuity and large financial losses. By examining security data, Machine Learning technology is essential for providing an automated, dynamically improved, and up to security system. Large datasets may be handled by Machine Learning technology, which also provides greater visualization capabilities. Due to built-in weaknesses and threats that can be used at any stage in the system, supply chain security is difficult. Our study's findings demonstrate that the majority of rules exhibit steady support and confidence values, enabling the prediction of Cyber Attacks over a period of days without the need for daily rule updates.

Keywords: Intrusion Detection, Support Vector Machine, Decision Tree, Deep Neural Networks, Logistic Regression

Authors

Dr. S. Raja Ratna
Assistant Professor
Department of Computer Science Engineering
SRM Institute of Science and Technology, Chennai, India

Dr. G. Gangadevi
Assistant Professor
Department of Computer Science Engineering
SRM Institute of Science and Technology, Chennai, India

Dr. J. Jospin Jeya
Assistant Professor
Department of Computer Science Engineering
SRM Institute of Science and Technology, Chennai, India

I. INTRODUCTION

Data scientists use a variety of interdisciplinary skills, including programming, mathematics, and statistics. Data scientists can find hidden patterns in unstructured data that can help guide important business decisions by utilizing tools, algorithms, and Machine Learning approaches.

Since sensitive data travels more and more into the digital sphere, cyber security has grown to be a critical issue in the current era. Data, computer networks, and other digital assets belonging to a company must be protected from invasions as cyber-attacks increase in frequency and severity. In order to do this, harmful attacks must be prevented on hardware, software, database systems, and related infrastructures.

Attackers utilize a variety of techniques, including viruses, malware, Trojan horses, logic bombs, and backdoors, to compromise systems by taking advantage of known software flaws.

Many systems and applications have been created employing sophisticated detection techniques to identify and stop Cyber Attacks. These technologies are dependent on an Industrial Control System (ICS), which is essential for monitoring and controlling vital resources.

II. LITERATURE SURVEY

Numerous kinds of study for intrusion detection systems have developed, but the emergence of Big Data has complicated conventional methods to dealing with significant quantities of data. As a result, a growing number of researchers are investigating Big Data methods, especially Machine Learning techniques, in order to develop efficient and quick breach detection systems. In this part, we highlight some academics that have used Machine Learning and Big Data methods for detection to address Big Data problems. Standard methods for intrusion detection are becoming increasingly difficult to apply as the field of Data Science evolves. As a result, a lot of researchers want to develop an accurate and quick intrusion detection system using Data Science techniques. In this section, we provide various instances of researchers that handled Data Science by utilizing algorithms based on machine learning for intrusion detection.

Rahman. [1] The writers used a data mining method called the J48 decision tree algorithm to mitigate cyber-attacks. They extracted historical trends from earlier cyber-attack cases and created a predictive algorithm to determine future occurrences of such assaults. The suggested approach was evaluated using openly accessible Canadian Institute of Cyber security datasets. The predictive model was effective in detecting DDoS, PortScan, Bot, Brute force, SQL Injection, and Heartbleed cyber assaults.

T. Gopalakrishnan [2] used BMO-DBN, a technique that uses the barnacles mating optimizer (BMO) to optimise deep belief networks (DBNs) for detecting cyber-attacks in mobile edge computing (MEC) settings, in their study. The suggested model has three main components: traffic forecasting, data unloading, and assault detection. The researchers used bidirectional long short term memory (BiLSTM) to estimate traffic first to enable efficient data offloading.

In their research, Dr. K. Jayasakthi Velmurugan [3] suggested a machine learning technique that can reach high levels of accuracy by incorporating entropy, precision, recall, F1 score, sensitivity, and specificity metrics in their research. The study concentrated on four different assault types: DOS, R2L, U2R, and Probe are the four kinds of assaults. The writers also proposed future improvements, such as the automation of outcome predictions via a desktop or online programmer, to maximize the efficiency of artificial intelligence (AI) deployment.

Amerasinghe and colleagues [4] introduced the SARIMA (Seasonal Auto Regressive Integrated Moving Average), a technique of time sequence as an extension of the ARIMA model for prediction in their study. Time series analysis, which includes seasonal-level moving average modelling, auto-regression, and differencing, was used by the writers. Based on a large database, they used data mining methods to identify, prevent, and predict cyber-attacks. The system contains a trained model that, using a specialized process, halts previously recognized assaults.

Huan Long [5] suggested using an improved iterative algorithm an interval DSSE-based approach to identify cyber-attacks on FIDA (IIA). The method entails detecting cyber-attacks when The traditional DSSE model estimates a state variable that exceeds the pertinent range specified by the interval DSSE model. Extensive testing was done to fully assess the cyber-attack model and detection technique, and the IIA algorithm outperformed the MC and IGE algorithms. Future study will concentrate on developing methods to minimize expected state delay and showing an improved three-stage DSSE architecture that can improve state prediction accuracy.

Celestine-iwendi.[6]suggested key plit Watermark, a watermarking-based method, to protect software code from online threats. The programme identifies buzzwords in the code, splits it based on the keyword, and generates a unique key. The method is dependable and safe, with only a small amount of processing overhead. Watermarks are not required as input for the extraction and embedding methods and the code is not altered in any way, rendering it opaque. In future study, the method will be modified and evaluated for its suitability with application-specific software written in various computer languages, each of which may contain a different set and quantity of keywords.

Sparjan's suggested method [7] employs numerous algorithms to learn from a prior dataset for better performance comparisons. A model is built based on this dataset, and success measures are computed and contrasted. In this suggested model, the KDD Cup99 dataset was used.

Fahima Hossain [8] discovered Random Forest and Gradient Boosting Machine Learning methods produced the highest results among the classifiers in her study. As a consequence, these techniques were chosen for inclusion in the model. The research included the use of five standard datasets, includingthe NSL-KDD dataset.

Chris Few [9], an independent freelancer, used SecuriCAD software to build a cyber-security model. SecuriCAD is a utility for creating system models from a cyber-security standpoint, and it evolved from an older tool known as CySeMoL. Several object kinds, including firewalls, clients, hosts, networks, protocols, services, and access restrictions, were used in the model's construction over time.

Bilen [10] has created a system that uses Machine Learning to analyze two distinct cybercrime models and forecast the effect of particular features on determining the method and perpetrator of the assault. The system uses machine learning techniques, including SVM, RF, Logistic Regression, XGBoost, DT, KNN, and NB. The original model attempts to forecast the character of the attack by gathering information on various aspects such as the crime, gender, age, occupation, finances, marital status, educational background, harm caused and perpetrator. The other model aims to identify the perpetrator of the crime by considering factors. The model predicts prospective victims' traits as well as the kinds of attacks they may encounter.

Al-Abassi [11] recommended utilizing deep learning to recognise Cyber Attacks in Industrial Control Systems (ICS). To balance the imbalanced original data and generate new representations, the proposed technique uses a deep representation-learning algorithm. To detect Cyber Attacks, an ensemble deep learning technique is used, which employs DNN and DT classifiers. The suggested system makes use of databases from the gas pipeline (GP) and secure water treatment (SWT), which include both regular and attacksamples. Once an attack has been discovered, it is essential to determine its nature and position in order to keep computational efficiency. Future study will concentrate on improving the suggested method's accuracy and creating an extra model to identify various kinds of assaults and their locations. The objective of Abel Yeboah-study Ofori's article [12] is to improve the security of the online supply chain by examiningand anticipating risks. To achieve their goal, they combined Machine Learning (ML) methods with cyber threat intelligence (CTI). To show the efficacy of their strategy, they gathered CTI data and used different ML algorithms to generate predictive analytics. The Majority Polling method was used to generate a summary of potential hazards from all of the ML algorithms. The sample in their research was the Microsoft Malware Prediction collection.

Sango Doyin [13] used GNB, QDA, KNN, and CART, four popular supervised learning techniques, to identify and categories DDoS attacks that cause flooding. They intend to find solutions by identifying and classifying DDoS flooding attacks on SDNs using low-cost machine learning (ML) techniques. The KDD-99 and NSL-KDD databases, which comprise three types of DDoS flooding attacks such HTTP, TCP, and UDP, were used for experimentation. These protocols were chosen for flooding attacks because they handle a significant amount of online application traffic. The researchers used Mininet to simulate DDoS flooding assaults on the SDN architecture (LOIC). Future work will concentrate on improving the suggested method's accuracy and creating an extra model to recognize various kinds of assaults and their locations.

The authors of this research [14] used several Machine Learning algorithms to evaluate datasets and forecast the existence of CSC nodes, including LR, SVM, DT, and MV. An illustration from Microsoft's Malware Prediction website was used to illustrate the efficacy of their strategy. For the training data, the conclusions of the Logistic Regression, Decision Tree, and SVM algorithms were combined using Majority Counting. The data bits were processed and linked using a workflow. The accuracy, parameter estimation, and predictions were evaluated using 10-fold cross-validation and AUC ROC curve analysis. The Decision Tree algorithm has been shown to be successful in finding and forecasting patterns in Cyber Attacks in cyber supply chain predictive analytics.

The authors [15] have created a productive and flexible IDS using deep learning. The output layer is in charge of changing the output format to the necessary format. Data categorization is used during network training to identify network messages as either normal (0) or malicious assault (1), and numerical values such as real values are included in data values. Age and other categorical statistics can also be provided. The technique can also identify Denial of Service assaults and give time stamps to fields to determine when packets are typically received.

III. DEEP NEURAL NETWORK AND DECISION TREE CLASSIFIER FOR IDENTIFYING CYBER ATTACKS

There have been several studies on Cyber Attack systems introduced. Traditional methods for processing big data have gotten more complicated with the rise of big data. As a result, several academics are working to create quick and precise Cyber Attack systems using big data techniques. To evaluate if his network traffic was an attack or not, the author employed the k-means algorithm in Spark's Machine Learning package. In the proposed strategy, training and assessment are conducted using the KDD Cup 1999. This suggested strategy by employing feature selection techniques.

Recommendation, the researchers present the suggested methodology as well as the tools and methods used. The diagram below displays the Spark Chi SVM algorithm. The steps of the proposed version can be summarized as follows. Fresh expressions are infused throughout the ensemble. An attack detection model with learning capabilities developed specifically for ICS setups.

For the purpose of detecting Cyber Attacks, this model uses a Deep Neural Network (DNN) and a Decision Tree (DT) classifier. Based on the results of 10-fold-cross-validation using genuine ICS datasets, the effectiveness of the recently published version is evaluated and shown to outperform the standard method. For this evaluation, classifiers including AdaBoost, DNN, and Random Forest were used (RF).

1. **Collection of Data:** Raw data is frequently gathered from big populations of individuals or sensors that collect data on an item or event. This information can be used in conjunction with a particular kind of algorithm to allow computers to forecast future occurrences and make judgments without human intervention. Yet, gathering data is just the first stage in the machine learning model training process. How successfully machine learning algorithms predict outcomes are significantly influenced by the caliber and applicability of the training data set. Working with raw data might result in a number of problems in real life.

   For instance, the data gathered could not be pertinent to the problem statement, which would make it less valuable for developing precise models. Also, it's possible that some forecasts may contain missing data, such as empty columns or absent images. Another problem is that the data may have some groups or categories that are underrepresented, which might cause the model to perform poorly. To ensure that raw data is suitable for use in Machine Learning models, it is crucial to carefully gather and pre-process it.

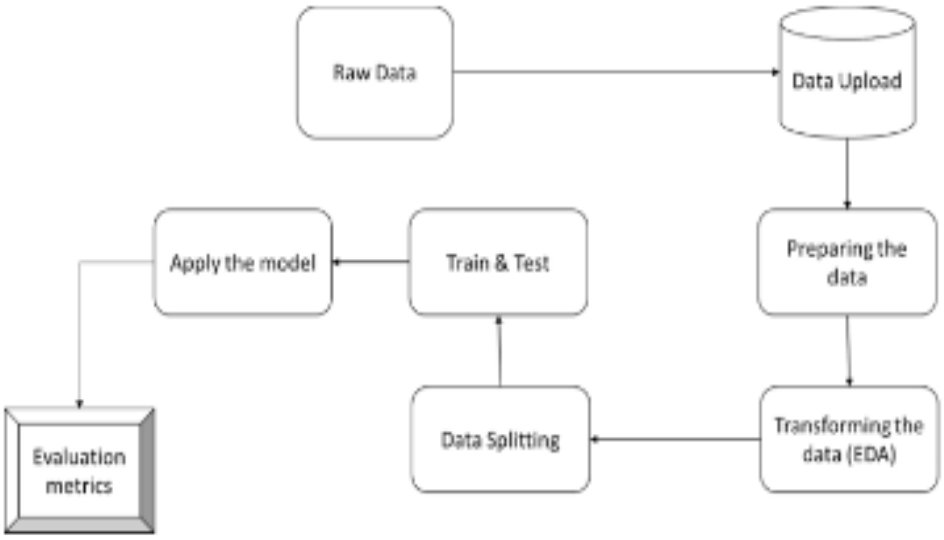


Figure 1: System Architectural Design

2. **Pre-Processing the Data:** Inconsistencies, fragmentation, and possible incompleteness are common in data gathered from the actual world when it comes to certain activities or patterns. Data mining methods could not properly find patterns in this complex data, so using them would not produce excellent results. The quality of the data as a whole must be improved more than anything else. In order to enhance the precision and quality of results, pre- processing data is a crucial stage in Machine Learning. It entails converting raw data into a more manageable format. The purging of data, which includes eliminating replicas, fixing inconsistencies, and imputed values that are missing is a crucial part of pre-processing. Data clarity is improved and noise is reduced as a result.

   Feature scaling is an additional consideration that attempts to uniformly distribute data throughout every aspect in order to prevent any one feature from taking over the model.

   Additionally, feature selection is a key method for removing superfluous features and simplifying the model, both of which can enhance model efficiency. Additionally, data pre-processing includes converting categorical factors into numerical values and creating new features using feature engineering methods.

3. **Extraction of Features:** The next stage is to reduce the amount of attributes in the data using a technique called feature extraction. In contrast to feature selection, which rates the pre-existing qualities according to how valuable they are for prediction, feature extraction includes changing the attributes themselves. To create the modified attributes, or features, the original attributes are linearly concatenated. Lastly, we train our models using the classifier technique. We use the collected, labeled dataset. The remaining labeled data that we have will be used to evaluate the models.

4. **Evaluating the Model:** Finding the model that reliably predicts future outcomes and depicts our data is advantageous. The use of training data to assess model performance is not permitted in Data Science, as this can result in the development of too optimistic and fitting models. In Data Science; there are two common approaches for evaluating models: hold-out and cross-validation. By assessing the model's performance on a different test set thatwas not utilized in training, both techniques try to avoid over fitting. The average performance over several runs is often used to evaluate the effectiveness of each categorization model. The outcomes can then be presented in the chosen format, frequently as a graph that illustrates thecategorization of the data.

5. **Data Splitting and Testing:** Data science requires the separation of a dataset into various subsets, or data splitting. Data used to train and evaluate a model, including testing data. Data are often partitioned at random to ensure that each subset appropriately represents the overall dataset. This process is essential for developing reliable and accurate machine learning models that are basedon data.

   The model's performance on fresh, untested data is then assessed using the testing set. This enables us to assess a model's ability to generalize to novel contexts and to gauge effectiveness of several models.

6. **Training the Model:** In order to train the model, examples of inputs and their associated outputs are provided, and the model's parameters are adjusted to reduce the discrepancy between expected and actual outputs. The training set should be substantial enough to allow the model to generalize to fresh, unexplored data, but not sizable enough to cause excessive training time or model complexity.

   Testing set is used to test the model once it has been trained, evaluated for efficacy, and tested. To accurately predict the target variable from unobserved data and minimize a preset loss function, the model's parameters must be changed. Building a machine learning system that works well requires careful evaluation of the data, model design, and optimization method. Training a model is a crucial phase in the process.

7. **Exploratory Data Analysis:** It is the process of analyzing data to spot patterns, trends, and potential issues. Data visualization, summary statistics, data cleansing, correlation analysis, and dimensionality reduction are some of the approaches used to better comprehend the data and spot any potential problems that can have an impacton further study. EDA is a crucial first stage in any project involving data analysis because it ensures that the data is thoroughly understood and that any potential problems or restrictions are found before more in-depth analysis is carriedout.

   The effectiveness of a machine learning model is evaluated using a range of measures, a few of which could change based on the job at hand and the model being used. For classification difficulties, metrics like accuracy, precision, recall, and F1-score are typically utilized, While problems with regression can be addressed using mean squared error, mean absolute error, and R-squared. The correct measures must be chosen, and the trade-offs between different metrics must be considered, to ensure that the model answers the current challenge. The evaluation matrix is a crucial tool for assessing the model's efficacy and identifying areas that could use improvement.

8. **UML Diagram:** The use case diagram is an illustration of how users or other external systems interact with a software programmer or system to show its capabilities and requirements. The actor, use case, and relationship diagram shows the system from the viewpoint of the user. It is a useful tool for system design and development, stakeholder communication, and ensuring that the system satisfies user demands. In general, use case diagrams are essential for understanding how a software application or system will really be used.

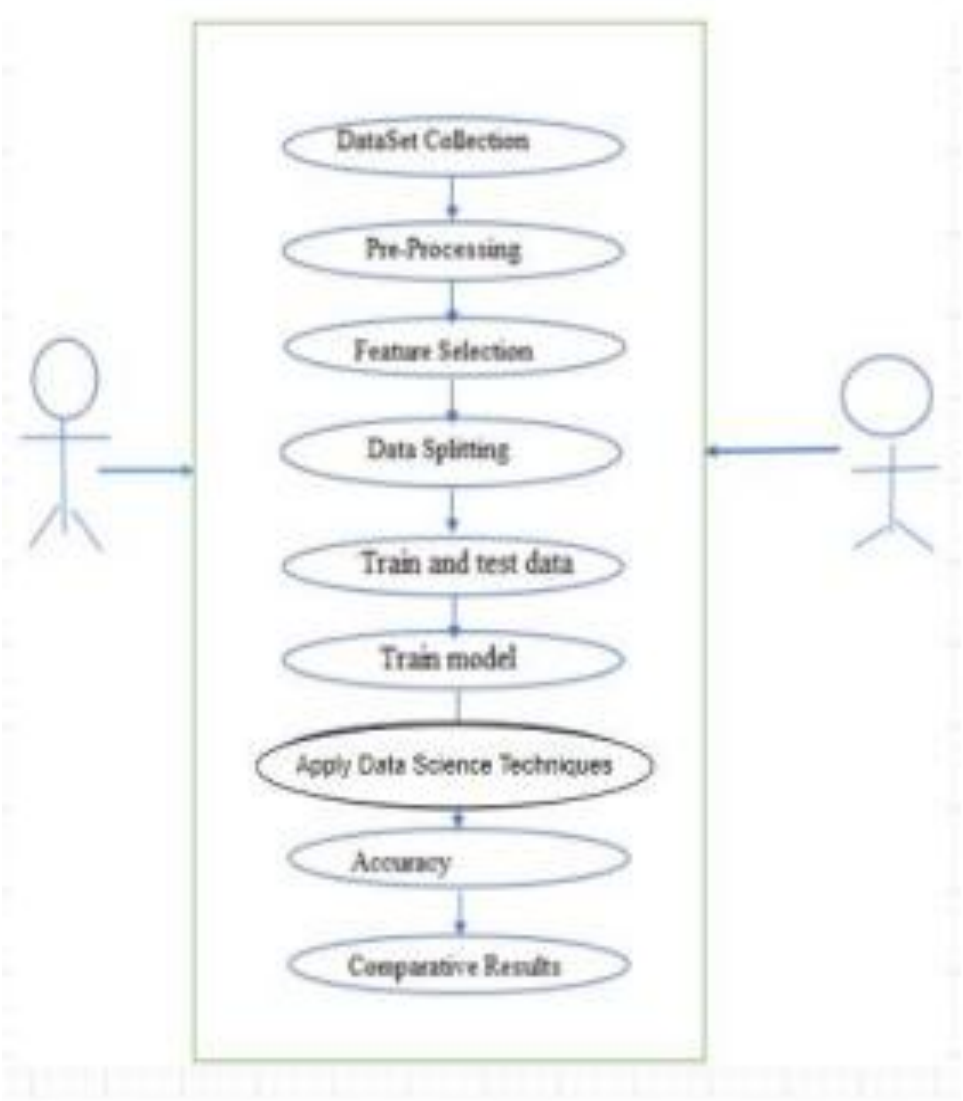


Figure 2: UML Diagram

9. **Sequence Diagram:** A sequence diagram shows the messages that are sent between objects and their lifelines to show how dynamically a system or process behaves. It replicates how things interact, allowing for the detection of design defects and inefficiencies and the subsequent implementation of necessary improvements.

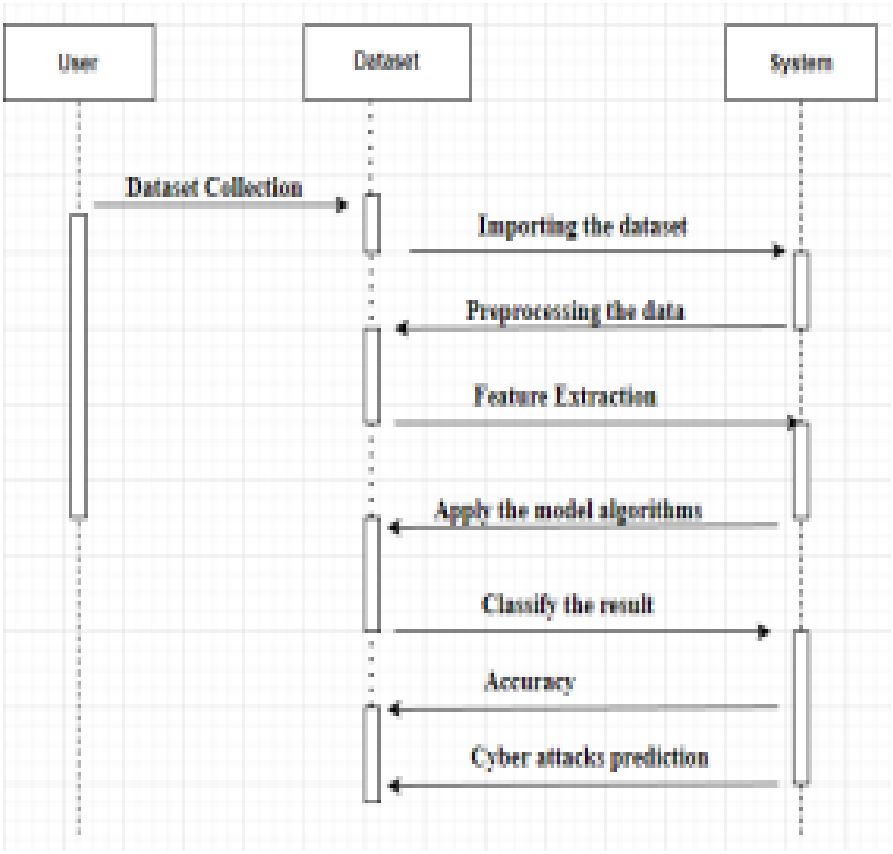


Figure 3: Sequence Diagram

For system design and development, sequence diagrams are useful because they provide a clear, concise image of how various system or process components interact with oneanother.

IV. ALGORITHMS USED

1. **Deep Neural Networks (DNN):** A deep neural network is modeled after the way the human brain works and is organized. The model includes of layers of nodes, or neurons that are linked and work together to process and make sense of input data. A DNN's architecture might change significantly based on the particular task and dataset that it is intended to handle. A variety of activities can be performed using deep neural networks. They are an effective method for resolving challenging Machine Learning issues and have produced state-of-the-art outcomes in numerous disciplines.

2. **Decision Tree (DT):** Both classification and regression problems can be accomplished using Machine Learning method known as decision trees. The data are repeatedly divided into subsets according to the greatest distinguishing attributes. The decision tree procedure will still result in a branch that is only a small portion of a much bigger tree. Yet, this algorithm cannot be disregarded regardless how straightforward it is. It offers clear information about feature importance, and by looking at the tree's structure, it is simple to understand the links between the characteristics. With the ability to handle both categorical and numerical data, decision trees have the advantage of being simple to

comprehend and analyze. They can, however, over fit and be sensitive to minute changes in the data. To enhance the performance of decision trees, it is crucial to employ strategies like pruning and ensemble.

3. **Random Forest (RF):** It is a learning technique that combines information from different decision trees to make predictions. The fundamental concept is to construct many decision trees, aggregate their predictions, and then achieve a more precise and reliable prediction. In comparison to decision trees, random forests are more accurate and resilient to noise and outliers. Building, many trees, though, can be computationally expensive and the model might be challenging to understand. Applications including classification, regression, and feature selection frequently make use of random forests.

4. **Support Vector Algorithm (SVM):** Popular supervised learning algorithms for categorization and regression analysis include Support Vector Machines (SVMs). They are especially helpful for projects like bioinformatics, image identification, and text classification. SVMs operate by finding a hyper plane in a high-dimensional space that can distinguish between various data groups. The support vectors, or closest data points for each class, are used to determine the hyper plane by maximizing the gap between them. SVMs can manage both linearly separable and non- linearly separable data by utilizing a kernel function. In comparison to other machine learning methods, SVMs have a number of benefits. They work well in high-dimensional environments, consume little memory, are adaptable, and are less prone to over fitting. For big datasets, they can be computationally demanding and sensitive to the regularization parameter and kernel function selection.

5. **Logistic Regression (LR):** Logistic regression is frequently used in classification problems, such as predicting the category or class of new observations. In accordance with the values of the input variables, the model forecasts the likelihood of a specific result. In logistic regression, the dependant variable has binary values, often 0 or 1. In logistic regression, the independent factors may be either continuous or categorical. Based on the values of these independent factors, The dependent variable will most likely be 1, according to the model's prediction. To do this, it applies a particular mathematical function known as the logistic or sigmoid function to the output of a linear regression model, converting it into a chance value between 0 and 1. The chance of the dependent variable being 1 can then be deduced from this probability value. The simplicity, interpretability, and simplicity of application of logistic regression make it superior to other classification algorithms in many ways. Predicting the probability of binary outcomes, like whether a patient will contract a disease or whether a customer will churn, is widely used in a variety of disciplines, including medicine, finance, and social sciences.

V. RESULTS AND IMPLEMENTATION

1. **Integration Testing:** Software unification experimentation is the gradual integration of two or more linked operating system components on a single application. The goal is to produce degradations brought on by interface flaws. The highest-level module is often tested first, followed by the lower-level modules or components it interacts with. This method is useful for identifying issues with the high-level architecture of the system.

According to this method, the higher-level modules that rely on the lower-level modules are tested afterward. It is helpful for figuring out problems with how the different parts are implemented. This strategy entails assembling and testing each system module or component separately. It is helpful for detecting problems with the overall architecture of the system and the interactions between the parts.

2. **Functional Testing:** Functional testing offers organized proof that the functionalities under test are present and operating properly in compliance with the technical and business requirements, system documentation, and user guides. Software testing includes functional testing, which is crucial because it concentrates on confirming the system's functioning and behavior in various scenarios. Functional testing's main objective is to make sure the programme operates as intended and complies with all necessary specifications. By examining how software behaves and functions in various scenarios, functional testing is essential for guaranteeing the quality and dependability of software systems.

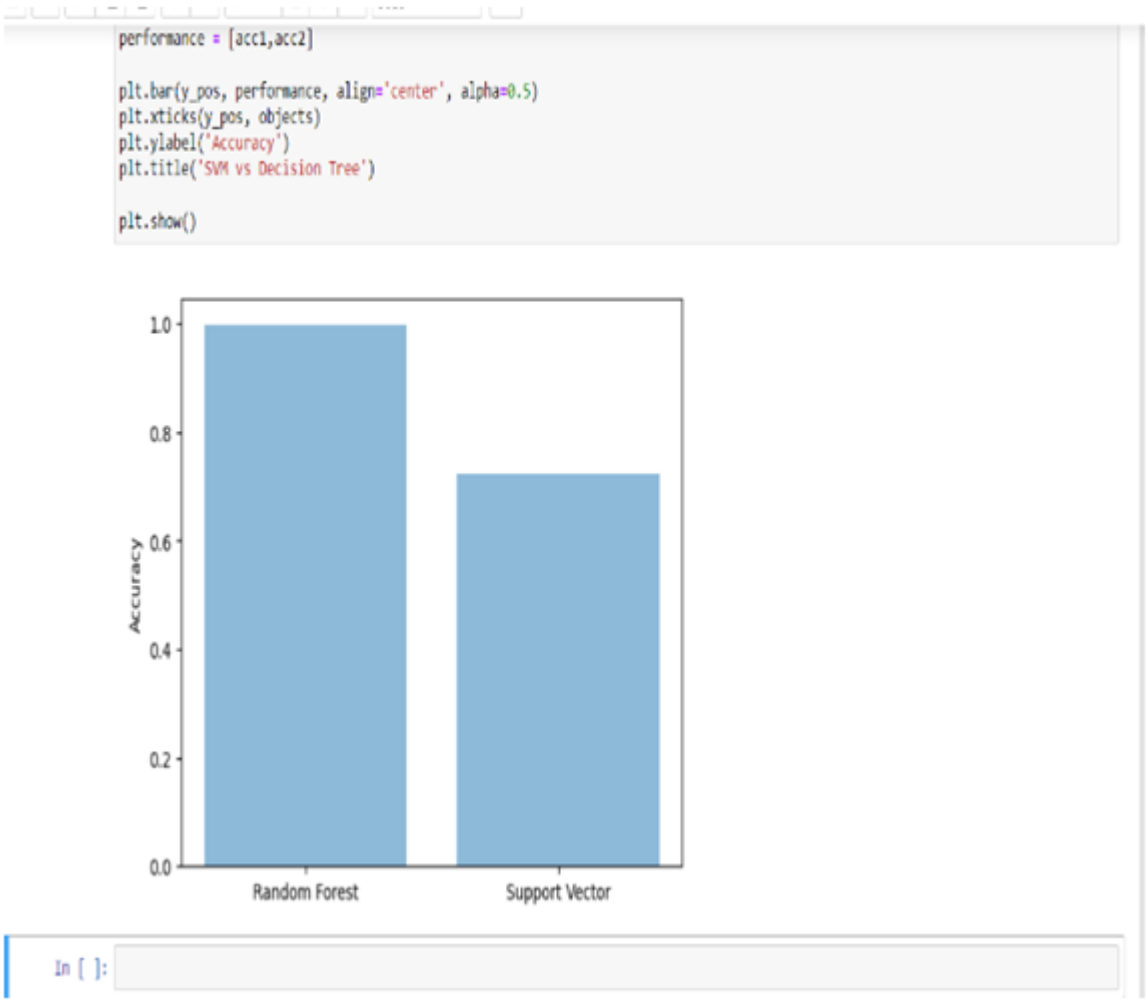
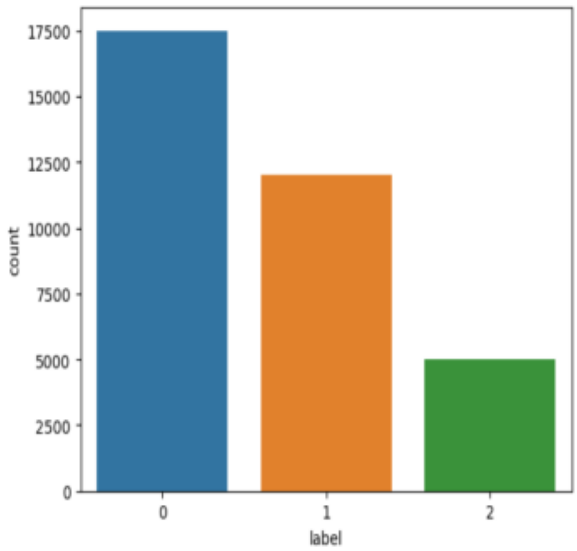


Figure 4: The value count of label in data set**.**

```
In [41]: sns.countplot(x='label',data=newdata1)
Out[41]: <AxesSubplot:xlabel='label', ylabel='count'>
```




```
In [27]: x = newdata1.iloc[:,newdata1.columns!='label']
         y = newdata1.iloc[:,newdata1.columns =='label']
```

Figure 5: Accuracy score graph between Random forest andsupport vector

```
objects = ('chi svm','chilogistic regression')
y_pos = np.arange(len(objects))
performance = [acc1,acc2]

plt.bar(y_pos, performance, align='center', alpha=0.5)
plt.xticks(y_pos, objects)
plt.ylabel('Accuracy')
plt.title('SVM vs Decision Tree')

plt.show()
```

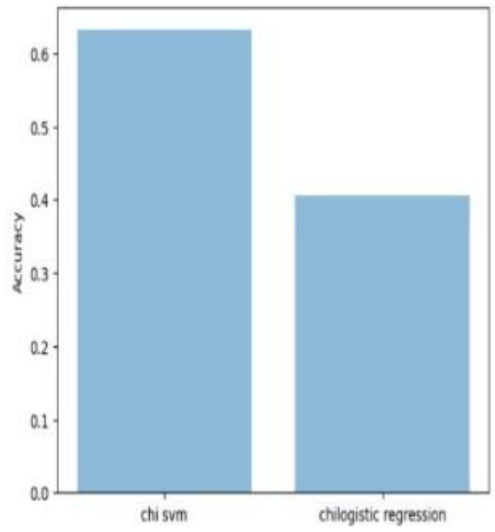


Figure 6: Accuracy score graph between chi svm and chi logistic regression

VI. CONCLUSION

The suggested technique incorporates a new architecture, a balanced representation of an unbalanced raw dataset, and a representative deep learning model. The additional phrases are then employed in complex sentences. Identify network attacks using learning methods based on DNN and DT Two independent datasets of Industrial Control Systems (ICS) were used, both derived from actual data-critical infrastructure, to assess the effectiveness of the suggested approach. The higher f1 score of the recommended technique showed that it performed 10% better than the traditional classifier. It achieves proper evaluation and generation of data in both datasets, with 95.86% and 99.67% accuracy on the dataset. Conventional classification methods like RF and DNN are the results.

REFERENCES

[1] Md Anisur Rahman, Yeslam Al-Saggaf, Tanveer Zia, "A Data Mining Framework to Predict Cyber Attack for Cyber Security", 15th IEEE Conference on Industrial Electronics and Applications (ICIEA) ,IEEE Access pp.207-212,2020.

[2] T. Gopalakrishnan , D. Ruby , Fadi Al-Turjman , Deepak Gupta ,Irina V. Pustokhina , Denis A. Pustokhin , And K. Shankar , "Deep Learning Enabled Data Offloading With Cyber Attack Detection Model in Mobile Edge Computing Systems", IEEE Access, vol. 8, pp. 185938-185947,2020.

[3] Dr.K.Jayasakthi Velmurugan, R.Rajasutha, S.Swetha, "Prediction Of Cyber Attack Using big data",IRJET,vol. 09,pp. 858-861, 2022.

[4] A.M.S.N. Amarasinghe, W.A.C.H. Wijesinghe, D.L.A. Nirmana, Anuradha Jayakody, A.M.S. Priyankara. "AI Based Cyber Threats and Vulnerability Detection, prevention and Prediction System", International Conference on Advancements in Computing (ICAC), pp. 363-367,2019.

[5] Huan Long, Member, IEEE, Zhi Wu, Member, IEEE, Chen Fang, Senior Member, Xinchi Wei, Wei Gu, and Huiyu Zhan, "Cyber-attack Detection Strategy Based on Distribution System State Estimation", JMPSCE, Vol. 08, No. 4, pp. 669-677, 2020.

[6] Zunera Jalil , Abdul Rehman Javed, Rajesh Kaluri , Thippa Reddy G ,Gautam Srivastava , Celestine Iwendi And Ohyun Jo, "KeySplitWatermark: Zero Watermarking Algorithm for Software Protection Against Cyber-Attacks", IEEE Access, Vol. 8, pp. 72650-72659,2020.

[7] Sparjan S , Deepan Raj M , Suriya prakash T , Senthil K, Preetha M. "Prediction Of Cyber-Attacks Using Data Science Technique",IJARIIE,Vol. 8, pp. 4239-4246, 2022.

[8] Fahima Hossain, Marzana Akter and Mohammed Nasir Uddin, "Cyber Attack Detection Model (CADM) Based on Machine Learning Approach", ICREST, pp. 567-572, 2021.

[9] Chris Few, James Thompson, Kenny Awuson-David, "A case study in the use of attack graphs for predicting the security of cyber-physical systems", IEEE Explore,2021.

[10] Abdulkadir Bilen, Ahmet Bedri Özer, "Cyber-attack method and perpetrator prediction using Machine Learning algorithms", PeerJ Comput. Sci.,2021.

[11] Abdulrahman Al-Abassi , Hadis Karimipour, Ali Dehghantanha , , And Reza M. Parizi , "An Ensemble Deep Learning-Based Cyber-Attack Detection in Industrial Control System", IEEE Access, Vol. 8, pp.83965-83972, 2020.

[12] Abel Yeboah-Ofori , Shareeful Islam , Sin Wee Lee , Zia Ush Shamszaman , Khan Muhammad , Meteb Altaf , And Mabrook S. Al-Rakhami, "Cyber Threat Predictive Analytics for Improving Cyber Supply Chain Security", IEEE Access, Vol. 9, pp. 94318-94335, 2021.

[13] Abimbola O. Sangodoyin , Mobayode O. Akinsolu , Prashant Pillai , And Vic Grout, "Detection and Classification of DDoS Flooding Attacks on SoftwareDefined Networks: A Case Study for the Application of Machine Learning", IEEE Access, Vol. 9,pp.122495- 122508, 2021.

[14] Abel Yeboah-Ofori , Charles Boachie , "Malware Attack Predictive Analytics in a Cyber Supply Chain Context Using Machine Learning", IEEE-ICSIoT,2019.

[15] Ahmad Hijazi, El Abed El Safadi, Jean-Marie Flaus. "A Deep Learning Approach for Intrusion Detection System in Industry Network",Researchgate,2019

