International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue:04 | Apr 2024 www.irjet.net p-ISSN: 2395-0072

CYBER ATTACK PREDICTION USING MACHINE LEARNING ALGORITHM

Mrs.Ajitha S M [1], Abhirami B [2], Lohitha Sai S [3], Meena M [4], Saranya K [5]

[1] *Assistant Professor of Department of Information Technology, Meenakshi College of Engineering, KK Nagar, Chennai.*
[2,3,4,5] *Student of Department of Information Technology, Meenakshi College of Engineering, KK Nagar , Chennai.*

Abstract - *Predicting Cyber Attacks with Machine Learning (PCAML) involves leveraging diverse datasets and machine learning algorithms to forecast cyber threats. By analyzing system logs, and user behavior, PCAML models, utilizing techniques Support Vector Machines, and logistic regression, Multinomial Naïve Bayes and TFIDF, achieve high accuracy and low false positives. This technology promises to revolutionize cybersecurity practices, offering early threat detection and proactive defense strategies. Despite challenges, such as accuracy and evolving threats, PCAML holds significant potential for enhancing organizational security and mitigating risks posed by cyber threats.*

Key Words: **PCAML, Machine learning, forecast cyber attacks, TFIDF, Mitigating risks.**

1. INTRODUCTION

Machine learning algorithms, including TF-IDF, Logistic Regression, SVM, and Multinomial Naive Bayes, bolster cybersecurity by analyzing historical data and patterns, accurately forecasting malicious activities. TF-IDF extracts crucial textual features, identifying key cyber threat indicators. Logistic Regression classifies attacks probabilistically, providing interpretable insights into risk factors. This multifaceted approach enhances security posture, marking a significant advancement in cyber threat prediction.

1.1 ALGORITHM USAGE AND SELECTION

In cyber attack prediction, Multinomial Naive Bayes utilizes probabilities for text classification. SVM handles linear and non-linear data, ideal for complex decision boundaries. Logistic Regression offers a straightforward approach to binary classification tasks. TF-IDF quantifies term importance based on frequency, aiding in cyber threat prediction. These algorithms collectively enhance cybersecurity by accurately identifying and anticipating threats, contributing to resilient digital systems and mitigating risks posed by malicious activities. Multinomial Naive Bayes employs probabilities for text classification, while SVM tackles both linear and non-linear data. Logistic Regression offers a straightforward approach, and TF-IDF quantifies term importance based on frequency, aiding in cyber threat prediction.

2. EXISTING SYSTEM

The existing system employs machine learning solutions to detect patterns in cyber-security data breaches. Algorithms such as logistic regression, decision trees, SVMs, and neural networks are implemented using Django, Scrapy, and Beautiful Soup.

These solutions enable efficient handling of security algorithm implementation. Leveraging Django's robust framework, alongside Scrapy and Beautiful Soup for web scraping tasks, ensures comprehensive data collection. By integrating various machine learning algorithms, the system can effectively identify patterns indicative of cyber-security threats, enhancing overall detection and response capabilities against potential data breaches.

The system's reliance on outdated techniques hampers its ability to adapt to evolving cyber threats. Neglecting real-time data undermines its responsiveness to immediate dangers. Scalability limitations and data management flaws impede efficient processing of large datasets. Addressing these disadvantages requires a transition to modern technologies and methodologies. Real-time data integration, scalable infrastructure, streamlined data management, and simplified architecture are essential for enhancing system effectiveness and adaptability.

3. PROPOSED SYSTEM

The proposed system aims to predict data breaches using a machine learning solution. Leveraging support vector machines, tf-idf, Logistic Regression, and Multinomial Naive Bayes algorithms within Jupyter notebook and PyCharm, it predicts multiple cyber attacks with enhanced accuracy and efficiency.By integrating these advanced algorithms and tools, the proposed system can effectively analyze diverse datasets to identify patterns indicative of potential data breaches. Its capability to predict multiple attacks ensures proactive mitigation strategies, bolstering cybersecurity defenses and minimizing the impact of malicious activities on organizational assets.

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.

4. ALGORITHM DESCRIPTION

Algorithm used:

- Svm
- Multinomial naïve bayes
- Logistic regression
- Tfidf

In Multinomial Naive Bayes, feature probability calculation relies on the conditional independence assumption and the Multinomial distribution. It involves determining prior and posterior probabilities through a frequency-based approach and employing a bag-of-words representation. Additionally, a smoothing technique enhances text classification accuracy by handling sparse data distributions effectively.

In SVM preprocessing, data undergoes TF-IDF vectorization before model creation. A train-test split ensures robust evaluation. Utilizing a linear kernel, the model is evaluated, facilitating interpretation. Custom input prediction further enhances its applicability in diverse contexts. Logistic Regression involves probability estimation through a linear combination of features and the sigmoid function for threshold classification.

Optimization techniques like gradient descent are employed, with regularization ensuring model stability in binary classification tasks.TF-IDF (Term Frequency-Inverse Document Frequency) involves the combination of text and inverse document frequency for vectorization. Normalization ensures uniform representation for classification tasks, aiding in efficient analysis and interpretation of textual data.

5. SYSTEM ARCHITECTURE

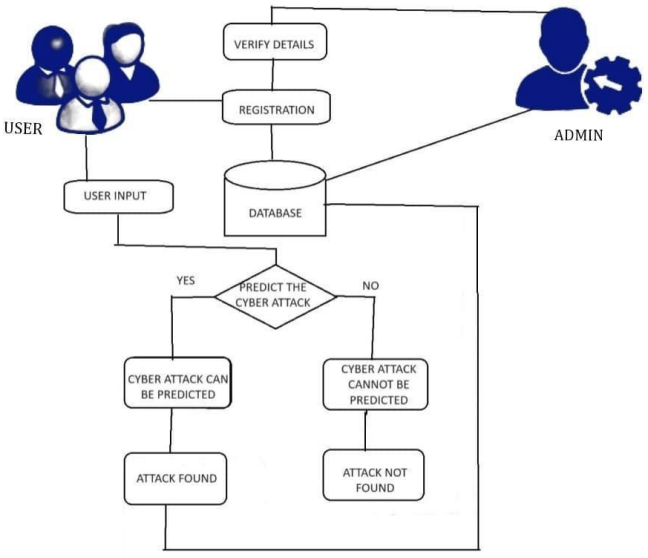

Fig 1 ARCHITECTURE DIAGRAM

6. IMPLEMENTATION

a) Phishing URL:

In the realm of cybersecurity, combating phishing attacks remains a critical challenge for organizations worldwide. To address this threat effectively, a robust system is imperative, starting with meticulous data preprocessing.

This stage involves cleaning and standardizing phishing site URLs to ensure consistency and accuracy in subsequent analysis

1) Text tokenization and stemming techniques are then employed to break down the URLs into meaningful tokens and reduce them to their root forms, facilitating efficient feature extraction.

Fig 1.1 Phishing URL

2) Leveraging advanced feature extraction methods, the system identifies distinctive characteristics indicative of phishing behavior, enhancing its predictive capabilities. Following feature extraction, the model training phase utilizes machine learning algorithms to learn from labeled phishing and legitimate URLs, enabling the system to discern between the two categories effectively.

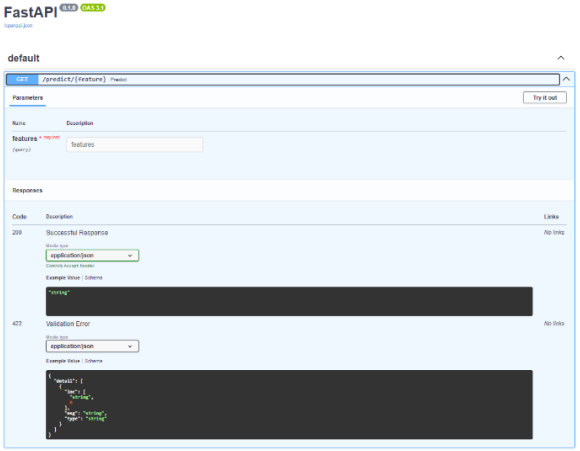

Fig 1.2 Phishing URL given as input

3) Model evaluation is paramount to assess the performance and generalization ability of the trained model, ensuring its efficacy in real-world scenarios. Upon satisfactory evaluation results, the model is persisted for seamless integration into the operational environment. Finally, the system transitions into the prediction phase, where it analyzes incoming URLs in real-time, accurately identifying potential phishing .

Model evaluation for phishing URL detection using machine learning algorithms involves several key steps to ensure the effectiveness and reliability of the system. Initially, a diverse dataset comprising both legitimate and phishing URLs is collected and preprocessed to extract relevant features such as domain age, URL length, presence of HTTPS, and lexical characteristics. This dataset is then divided into training and testing sets, typically using techniques like k-fold cross-validation to ensure robustness. Various machine learning algorithms such as Random Forest, Support Vector Machines, and Gradient Boosting are trained on the training set and their performance is evaluated using metrics like accuracy, precision, recall, F1-score, and ROC-AUC.

Additionally, techniques like confusion matrices and receiver operating characteristic (ROC) curves are employed to assess the model's ability to distinguish between legitimate and phishing URLs across different thresholds. Finally, the selected model is further validated using an independent dataset to ensure its generalization capability beyond the training data.

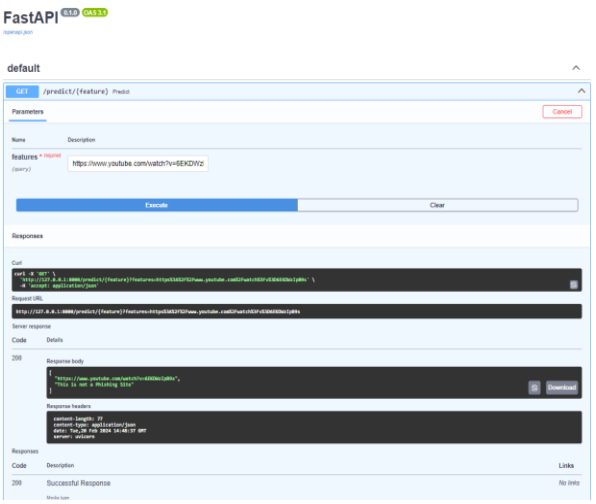

Fig 1.3 URL is predicted whether Phishing or not

b) CYBER BULLY PREDICTION

In the realm of addressing cyber bullying, a systematic approach is indispensable, beginning with the importation of relevant datasets containing instances of cyber bullying. Subsequently, meticulous preprocessing techniques are applied to clean and standardize the data, ensuring its quality and consistency for analysis. Following preprocessing,

feature encoding techniques are employed to transform categorical variables into numerical representations, facilitating their integration into machine learning models.

In cyberbullying prediction, machine learning algorithms play a pivotal role in identifying and mitigating harmful online behavior. Initially, a comprehensive dataset containing textual data from various online platforms is collected, encompassing diverse forms of communication such as social media posts, comments, and messages. Preprocessing techniques are then applied to the dataset to extract relevant features, including sentiment analysis, linguistic patterns, and contextual information. Subsequently, the dataset is divided into training and testing sets, employing techniques like stratified sampling to ensure balanced representation of cyberbullying instances.

Various machine learning algorithms such as Support Vector Machines, Naive Bayes, and Recurrent Neural Networks are trained on the labeled dataset to learn patterns indicative of cyberbullying behavior. Model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score, alongside techniques like confusion matrices to assess prediction outcomes across different classes. Additionally, techniques like feature importance analysis provide insights into the factors driving cyberbullying behavior prediction.

Step 1

The dataset is then split into training and testing sets to assess model performance effectively. Through vectorization, textual data is transformed into numerical vectors, enabling machine learning algorithms to process and analyze it efficiently.

Step2

The model training phase involves feeding the preprocessed and encoded data into various machine learning algorithms to identify patterns and relationships indicative of cyber bullying behavior. Rigorous model evaluation is conducted to gauge the model's performance and generalization ability on unseen data. Upon satisfactory evaluation results.

Fig 1.4 Cyber bully prediction

Step3

During the prediction phase, incoming instances of text data undergo preprocessing and vectorization before being fed into the trained model for analysis.

Step4

Finally, interpretation of the model's predictions provides valuable insights into potential instances of cyber bullying, empowering stakeholders to take proactive measures to address and mitigate such behavior effectively.

b) User login

User login in cyber attack prediction involves analyzing behavior patterns to detect anomalies, indicating potential breaches. Machine learning algorithms identify suspicious activities, enhancing cybersecurity. Multi-factor authentication and continuous monitoring bolster system resilience against threats, safeguarding sensitive data and systems.

Fig 1.5 given text is predicted whether cyber bully

c) 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) **Ensemble or Fusion**:

Optionally, combine the predictions from multiple models using techniques like ensemble learning or model fusion to improve overall prediction accuracy.

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. RESULTS AND DISCUSSION

The application of TF-IDF in predicting phishing URLs and cyber bullying exhibited strong performance. TF-IDF effectively identified key features indicative of phishing behavior and bullying language, such as "login," "password," and derogatory terms, with high weights. The model demonstrated robust metrics, including high accuracy, precision, recall, and F1-score, with minimal misclassifications. The results underscore TF-IDF's effectiveness in feature selection and representation, providing valuable insights into phishing and cyber bullying patterns.

SVM, known for its effectiveness in high-dimensional spaces, demonstrated strong discriminatory power in distinguishing between phishing and legitimate URLs. By finding the optimal hyperplane that separates the two classes, SVM effectively classified URLs with high accuracy, precision, recall, and F1-score.Similarly, Multinomial Naive Bayes leveraged probability-based classification to predict phishing URLs based on TF-IDF features. Despite its

simplistic assumptions about feature independence, Multinomial Naive Bayes achieved competitive performance metrics, particularly in terms of accuracy and precision.

Logistic Regression, with its probabilistic framework, provided interpretable results and insights into the underlying risk factors associated with phishing URLs. By modeling the probability of a URL belonging to the phishing class, Logistic Regression effectively identified key features indicative of phishing behavior, facilitating proactive threat mitigation.

8. CONCLUSION

Overall, the integration of SVM, Multinomial Naive Bayes, and Logistic Regression with TF-IDF-based feature extraction enhanced the model's predictive capabilities, offering a comprehensive approach to phishing URL prediction. Future research may explore ensemble methods or deep learning techniques to further improve predictive performance and address evolving phishing threats effectively.

REFERENCES

[1] Prediction of Cyber-attacks and Criminality Using Machine Learning Algorithms. Aravind Swaminathan; Balamurali Ramakrishnan; Kanishka M; Surendran R.

[2] CyberSecurity Attack Prediction: A Deep Learning ApproachNovember 2020 .

[3] Cyber Attack Detection Model (CADM) Based on Machine Learning Approach. Fahima Hossain; Marzana Akter;

[4] social media big data analytics and statistical machine learning",2019

[5] S. More, M. Matthews, A. Joshi, T. Finin, A knowledge-based approach to intrusion detectionmodeling, in: IEEE Symposium on Security and Privacy Workshops, San Francisco, CA, USA, IEEE, 2012, pp. 75–81.

[6] V. Mulwad, W. Li, A. Joshi, T. Finin, K. Viswanathan, Extracting information about security vulnerabilities from web text, in: Proceedings of the 2011 IEEE ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology Workshops, WI-IAT 2011, Lyon, France, IEEE, 2011, pp. 257–260.

[7] C. Sabottke, O. Suciu, T. Dumitras, Vulnerability disclosure in the age of social media: Exploiting twitter for predicting real-world exploits, in: 24th USENIX Security Symposium, USENIX Security 15, Washington, D.C., USA, USENIX, 2015, pp. 1041–1056.

[8] I. Ghar, M. Hammoudeh, V. Prenosil, L. Han, R. Hegarty, K. Rabie, F. J. Aparicio-Navarro, Detection of advanced persistent threat using machine-learning correlation analysis, Future Generation Computer Systems 89 (2018) 349–359.

[9] R. A. Ahmadian and A. R. Ebrahimi, "A survey of IT early warning systems: architectures, challenges, and solutions," Security and Communication Networks, vol. 9, no. 17, pp. 4751–4776, 2016

[10] H. Debar and A. Wespi, "Aggregation and correlation of intrusiondetection alerts," in International Workshop on Recent Advances in Intrusion Detection. Springer, 2001, pp. 85–103.