**PREDICTION OF CYBER ATTACKS IN REAL TIME**

**PROJECT REPORT**

***Submitted by***

**ABDUL RAHEEM A (201SE101)**

**SANJEEVKUMAR VS (201SE132)**

**AJAI T J (201SE103)**

***In partial fulfillment for the award of the degree***

# BACHELOR OF ENGINEERING

**in**

INFORMATION SCIENCE AND ENGINEERING



**BANNARIAMMANINSTITUTE OF TECHNOLOGY (An**

**Autonomous Institution Affiliated to Anna University, Chennai)**

**SATHYAMANGALAM-638401**

# ANNA UNIVERSITY: CHENNAI 600025

**OCTOBER 2023**

# BONAFIDE CERTIFICATE

Certified that this project report **“PREDICTION OF CYBER ATTACKS IN REAL TIME”** is the Bonafide work of **"ABDUL RAHEEM A (201SE101), SANJEEVKUMAR V S (201SE132), and AJAI T J (201SE103)"** who carried out the project work under my supervision.

|  |  |
| --- | --- |
| **Dr. J Nirmala Devi** | **Mr. S Satheesh Kumar** |
| **HEAD OF THE DEPARTMENT** | **ASSISTANT PROFESSOR** |
| Department of Information Science & Engineering | Department of Artificial Intelligence & Data Science |
| Bannari Amman Institute of Technology | Bannari Amman Institute of Technology |

**Submitted for Project Viva Voice examination held on……………**

**DECLARATION**

We affirm that the project work titled **“Prediction Of Cyber Attacks in Real Time”** being submitted in partial fulfillment for the award of the degree of **Bachelor of Engineering** in **Computer Science Engineering** is the record of original work done by us under the guidance of **Mr.Satheeshkumar S**, Assistant Professor**,** Department of Artificial Intelligence And Data Science. It has not formed a part of any other project work(s) submitted for the award of any degree or diploma, either in this or any other University.

|  |  |  |
| --- | --- | --- |
| **ABDUL RAHEEM A** | **SANJEEVKUMAR V S** | **AJAI TJ** |
| **(201SE101)** | **(201SE132)** | **(201SE103)** |

I certify that the declaration made above by the candidates is true.

**Mr SATHEESHKUMAR S**

**ACKNOWLEDGEMENT**

We would like to enunciate heartfelt thanks to our esteemed Chairman **Dr. S.V. Balasubramaniam**, Trustee **Dr. M. P. Vijayakumar**, and the respected Principal **Dr.C. Palanisamy for providing excellent facilities and support during the course of studyin this institute.**

We are grateful to **Dr. Nirmaladevi J**, **Head of the Department, Department of Information Science Engineering** for her valuable suggestions to carry out the project work successfully.

We wish to express our sincere thanks to Faculty guide **Mr. Satheeshkumar S, Assistant Professor**, **Department of Artificial Intelligence And Data Science,** for his constructive ideas, inspirations, encouragement, excellent guidance, and much needed technical support extended to complete our project work.

**We would like to thank our friends, faculty and non-teaching staff who have directly and indirectly contributed to the success of this project.**

**ABDUL RAHEEM A(201SE101)**

**SANJEEVKUMAR V S(201SE132)**

**AJAI TJ(201SE103)**

**INTERNAL EXAMINER EXTERNAL EXAMINER**

**ABSTRACT**

In the present hyper-associated world, the raising danger of digital assaults represents a significant gamble to people, associations, and countries. Anticipating and moderating these dangers has become principal. This venture dives into the domain of digital protection by saddling the force of information science and AI procedures to foresee and forestall digital assaults.

Key libraries like Pandas, NumPy, Seaborn, Matplotlib, Plotly, and Time are basic to our venture. They empower information control, representation, and transient investigation, giving basic bits of knowledge into digital assault designs.

AI models are at the core of our prescient framework. We utilize Calculated Relapse and Multinomial Credulous Bayes calculations, engaged by instruments like train\_test\_split for information parting and classification\_report and confusion\_matrix for execution assessment. Moreover, RegexpTokenizer and SnowballStemmer upgrade text preprocessing, while CountVectorizer works with highlight extraction. The pipeline is smoothed out utilizing make\_pipeline to improve model preparation.

Understanding assault vectors and weaknesses is fundamental, which is where Picture, Word Cloud, Beautiful Soup, Selenium, and NetworkX become possibly the most important factor. Beautiful Soup and Selenium help in web scratching for constant danger information, while NetworkX helps with dissecting network structures.

Moreover, to guarantee smooth sending and constant observing, we depend on uvicorn and fastapi for building a hearty Programming interface, working with connection with the prescient model. At long last, joblib guarantees model ingenuity for consistent combination into creation frameworks.

This project amalgamates these libraries and strategies, making a complete answer for foreseeing and forestalling digital assaults, at last bracing our computerized world's security foundation.

**Keywords:**

Cyber Attacks,Machine Learning,Predictive Analysis,Data Science,Security,FastAPI Deployment

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **S.NO** | **TITLE** | **PAGE NO** |
|  |  |  |
|  |  |  |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| **FIGURE NO** | **FIGURE NAME** | **PAGE NO** |
|  |  |  |
|  |  |  |

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| **ACRONYOM** | **ABBREVATION** |
|  |  |
|  |  |

**CHAPTER – 1**

**INTRODUCTION**

Network safety in the advanced time has turned into an irreplaceable aspect of our computerized lives. With the outstanding development of information and the rising inter-connectivity of frameworks, the danger of digital assaults increasingly poses a threat than at any other time. This presentation section makes way for our far reaching investigation of prescient examination in the domain of online protection.

Background of the work:

The consistently advancing scene of digital dangers requests proactive measures to protect delicate data and basic framework. Throughout the long term, digital assaults have filled in complexity and recurrence, making it basic to foster imaginative arrangements that can prudently distinguish and alleviate these dangers.

This work dives into the space of prescient examination to expect and counter digital assaults. By utilizing the force of information science and AI procedures, we expect to translate examples and irregularities in digital danger information, eventually strengthening our advanced protections.

Scope of the Proposed Work:

The consistently advancing scene of digital dangers requests proactive measures to protect delicate data and basic framework. Throughout the long term, digital assaults have filled in complexity and recurrence, making it basic to foster imaginative arrangements that can prudently distinguish and alleviate these dangers.

This work dives into the space of prescient examination to expect and counter digital assaults. By utilizing the force of information science and AI procedures, we expect to translate examples and irregularities in digital danger information, eventually strengthening our advanced protections.

ADVANTAGES :

SELENIUM :

Dynamic Web Scratching: Selenium succeeds in its capacity to associate with dynamic pages, which frequently contain basic ongoing data connected with digital dangers. Via computerizing web collaborations, it can actually scratch information from these pages, guaranteeing that investigators approach the most forward-thinking danger knowledge.

Cross-Program Similarity: Selenium upholds different internet browsers, making it flexible for scratching information from various sources. This similarity guarantees that network protection experts can get to danger information from different internet based stages without impediments.

Scriptable Errands: Selenium permits clients to prearrange different undertakings, for example, clicking buttons, finishing up structures, and exploring through sites. This usefulness is important while managing complex sites that expect collaboration to get to danger data.

Headless Perusing: Selenium can perform web scratching without the requirement for a graphical UI, a component known as headless perusing. This recoveries processing assets as well as empowers discrete information assortment, which is fundamental while checking vindictive sites or gatherings.

BEAUTIFUL SOUP :

HTML Parsing: Wonderful Soup succeeds in parsing HTML, going with it an optimal decision for removing organized information from website pages. With regards to network safety, this takes into consideration the extraction of important danger information from sites and gatherings.

Vigorous Label Search: Lovely Soup gives a straightforward and natural punctuation for looking and exploring through HTML reports. Network protection experts can without much of a stretch find explicit components on a website page, like danger pointers or catchphrases.

Information Cleaning: Wonderful Soup can clean and design scratched information, eliminating superfluous HTML labels and arranging issues. This is significant for setting up the information for additional investigation and representation.

Adjustable: Clients can tweak Lovely Soup's way of behaving by characterizing parsers and channels, taking into consideration custom-made web scratching arrangements. This adaptability is fundamental while managing a large number of sources and configurations in network safety.

FASTAPI :

High Performance: FastAPI is built on top of Starlette and Pydantic, offering exceptional performance in terms of request handling. In the context of cybersecurity, this means that the API can efficiently handle a large number of incoming requests, ensuring real-time threat data delivery.

Asynchronous Support: FastAPI supports asynchronous programming, allowing cybersecurity applications to handle concurrent tasks seamlessly. This is crucial for real-time monitoring and analysis of cyber threats as it ensures minimal latency.

Interactive Documentation: FastAPI generates interactive API documentation automatically. This feature is invaluable for cybersecurity professionals who need to understand and utilize the API's endpoints, making it easier to integrate the predictive cyber threat model into existing security infrastructure.

Security Features: FastAPI provides built-in security features, including input validation and authentication support, which are vital for securing access to the cyber threat prediction system. This ensures that only authorized personnel can access sensitive threat data.

Scalability: FastAPI's architecture is designed for scalability, allowing cybersecurity solutions to grow and adapt as the threat landscape evolves. This scalability is essential for organizations that need to handle an increasing volume of threat data and users over time.

**CHAPTER 2**

**LITERATURE SURVEY**

As of now, a large portion of the financial, business, social, social and legislative exercises and connections of nations, at all levels, including people, non-legislative associations and government and legislative organizations, are done in the internet. As of late, numerous privately owned businesses and government associations all over the planet are dealing with the issue of digital assaults and the risk of remote correspondence advancements.

1. Mr.S.S.Vasantha Raja , Aakash B, Avinash M , Gokul 4 , “Prediction of cyber attacks using Machine learning technique” 2022 *International journal of creative research thoughts (IJCRT)*,journal,Volume 10, Issue 6 June 2022,ISSN: 2320-2882.

The motivation behind this study was to the course of forecast examination is a course of utilizing a technique or innovation to investigate or animate a few obscure, unseen or confounded middle of the road processes in light of past and present states and afterward conjectured the outcomes. A different techniques required deduced information being a need and were challenging to segregate between typical burst deals and transition of dos assaults.

Cyber-attacks are quickly becoming one of the world's most serious problems. This crisis can be avoided by using real-time data to identify an attack and its perpetrator. The information can be obtained from the implementations of individuals who were subjected to cyber-attacks in forensic units. The information includes the criminal activity, the perpetrator's gender, impairment, and attack methods.

1. Aravind Swaminathan,Balamurali Ramakrishnan,Kanishka M,Surendran R, “*Prediction of Cyber-attacks and Criminality Using Machine Learning Algorithms*” 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT) , Sakheer, Bahrain,Print on Demand(PoD) ISBN:978-1-6654-5194-9 ,Print on Demand(PoD) ISSN: 2770-7458 , INSPEC Accession Number: 22475142 ,DOI: 10.1109/3ICT56508.2022.9990652

In this paper, research computerized misdeeds in three unmistakable models utilizing ML methods, and gauge the effect of the described traits on the spot of the electronic attack strategy and the culprit. In this examination, they utilize three ML calculations, Strategic Relapse, Arbitrary Woodland, and K-Closest Neighbor (KNN) and look at their adequacy in two unique models prior to finishing up with the model that has the best survivability for each sort of data record.

1. Hashim Albasheer ,Maheyzah Md Siraj ,Azath Mubarakali ,Omer Elsier Tayfour ,Sayeed Salih ,Mosab Hamdan ,Suleman Khan ,Anazida Zainall,"*Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques*",2022, 22(4), 1494; <https://doi.org/10.3390/s22041494>

An intricate and new instance of interruptions, new bugs, new security issues, and weaknesses are developing consistently. The assessment cycle of ready relationship research has similar difficulties of NIDS research, as AC is a reciprocal framework to NIDS. One of these difficulties is the inaccessibility of enough datasets. Nonetheless, various datasets have been utilized for testing Ready Connection explores, for example, KDD 99.

The thought is to empower security analysts to have instruments available to them that makes it simpler to convey the assault results with different partners who might have almost no network safety aptitude. Furthermore, with the proposed approach analysts mental burden can be diminished via naturally anticipating the results of assaults in the event that new goes after are found.

1. Prerit Datta,Natalie Lodinger,Akbar Siami Naminand Keith S. Jones , "*Predicting Consequences of Cyber-Attacks*",2020 IEEE International Conference on Big Data , 978-1-7281-6251-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/BigData50022.2020.9377825.

In this paper, they utilized AI models to anticipate the outcomes of digital assaults utilizing two famous word inserting techniques, that is to say, tf-idf and Doc2Vec. LinearSVCaccomplished the best execution for both the cases, which isreliable with past examination that showed LinearSVC to be appropriate for multiclass-order issues in normal language handling assignments.

This paper provides a survey of prediction, and forecasting methods used in cyber security. Four main tasks are discussed first, attack projection and intention recognition, in which there is a need to predict the next move or the intentions of the attacker, intrusion prediction, in which there is a need to predict upcoming cyber attacks, and network security situation forecasting, in which we project cybersecurity situation in the whole network.

5.Martin Husak, Jana Kom ´ arkov ´ a, Elias Bou-Harb, and Pavel ´ Celeda ,"Survey of Attack Projection, Prediction, and Forecasting in Cyber Security" , 2019 , COMST and y ERDF “*CyberSecurity,CyberCrime and Critical Information Infrastructures Center of Excellence*” (No.CZ.02.1.01/0.0/0.0/16 019/0000822).

Albeit many use cases were proposed, they can be decreased to a few primary use cases, in particular, assault projection and expectation acknowledgment, assault or interruption expectation, and security circumstance determining.