 

**DETECTING PHISHING WEBSITE USING MACHINE LEARNING**

### A PROJECT REPORT

***Submitted by***

|  |  |
| --- | --- |
| **GOWTHAM R** | **113019202012** |
| **JAGADESHWARAN S** | **113019205020** |
| **MANUEL MANASSEH G** | **113019205028** |
| **MARIMUTHU RAJA M** | **113019205029** |

*in partial fulfilment for the award of the degree of*

## BACHELOR OF TECHNOLOGY

IN

**INFORMATION TECHNOLOGY**

# VEL TECH HIGH TECH

**Dr. RANGARAJAN Dr. SAKUNTHALA ENGINEERING COLLEGE**

**An Autonomous Institution**

**MAY 2023**

# VEL TECH HIGH TECH

**Dr. RANGARAJAN Dr. SAKUNTHALA ENGINEERING COLLEGE**

**An Autonomous Institution**



## BONAFIDE CERTIFICATE

Certified that this project report **“DETECTING PHISHING WEBSITE USING MACHINE LEARNING”** is the bonafide work of **“GOWTHAM R (113019205012), JAGADESHWARAN R (113019205020), MANUEL MANASSEH G (113019205028), MARIMUTHU RAJA M (113019205029)”**

who carried out the

project work under my supervision.

**SIGNATURE SIGNATURE**

**Dr. M. MALLESWARI,** M.E, Ph.D. **Mrs. M.RAMYA,** M.E.

**HEAD OF THE DEPARTMENT SUPERVISOR**

ASSISTANT PROFESSOR

Department of Information Technology, Department of Information Technology, Vel Tech High Tech Dr. RANGARAJAN Vel Tech High Tech Dr. RANGARAJAN Dr. SAKUNTHALA ENGINEERING COLLEGE Dr. SAKUNTHALA ENGINEERING

COLLEGE

ii

## CERTIFICATE OF EVALUATION

College Name : VEL TECH HIGH TECH Dr. RANGARAJAN Dr. SAKUNTHALA ENGINEERING COLLEGE

Degree : BACHELOR OF TECHNOLOGY

Branch : INFORMATION TECHNOLOGY

Semester : VIII

|  |  |  |  |
| --- | --- | --- | --- |
| **S.**  **No.** | **Name of the Students** | **Title of the Project** | **Name, Designation & Department of the Supervisor** |
| 01 | GOWTHAM R | DETECTING PHISHING WEBSITE USING MACHINE LEARNING | Mrs. M. RAMYA ASSISTANT PROFESSOR  Department of Information Technology |
| 02 | JAGADESHWARAN R |
| 03 | MANUEL MANASSEH G |
| 04 | MARIMUTHU RAJA M |

### The report of the project work submitted by the above students in partial fulfilment for the award of degree, Bachelor of Technology/Engineering in

**for the viva voce examination held at Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College on has been evaluated and confirmed to be reports of the work done by the above students**

### INTERNAL EXAMINER EXTERNAL EXAMINER

***(Name, Designation and Signature with Date) (Name, Designation and Signature with Date)***

**iii**

### ACKNOWLEDGEMENT

We wish to express our sincere thanks to the people who extended their help during the course of our project work.

First of all we would like to express our deep gratitude to our beloved and respected

**FOUNDER & CHAIRMAN Col. Prof. Vel. Shri. Dr. R. RANGARAJAN B.E. (Elec.),**

**B.E (Mech.), M.S (Auto.), DSC, Foundress & VICE CHAIRMAN Dr. SAKUNTHALA RANGARAJAN M.B.B.S. and MANAGING TRUSTEE Mrs. MAHALAKSHMI KISHOREKUMAR, B.E., MBA(UK),Ph.D.,** for their kind encouragement and blessings. Our deepest gratitude and thanks to our Vice President M**r.K.V.D. KISHORE KUMAR, B.E., MBA (US).**

We also record our sincere thanks to our **Honorable Principal Dr.E.KAMALANABAN, B.E.,M.E., Ph.D.,** for his kind support to take up this project.

We are thankful and extremely grateful to our **Head of the Department Dr.M.Malleswari M.E.,Ph.D.** Computer Science and Engineering Department for her sustained help, guidance and inspiration in doing this project.

We would like to extend our sincere thanks to our **Project Coordinator Mr. Rajasekaran.R**

**.M.E,** for her continuous support for doing this project.

We would like to extend our sincere thanks to our **guide Mrs. M. Ramya, M.E,**for her constant technical support and stupendous encouragement, which enabled us to complete our project successfully.

Further the acknowledgement would be incomplete if we would not mention a word of thanks to our most **beloved PARENTS** whose continuous support and encouragement all the way rough the course has led us to pursue the degree and encouragement all the way through the course has led us to pursue the degree and confidently complete the project.

iv

## ABSTRACT

The Internet has become an indispensable part of our life; However, It also has provided opportunities to anonymously perform malicious activities like Phishing. Phishers try to deceive their victims by social engineering or creating mock up websites to steal information such as account ID, username, password from individuals and organizations. Although many methods have been proposed to detect phishing websites, Phishers have evolved their methods to escape from these detection methods.

One of the most successful methods for detecting these malicious activities is Machine Learning. This is because most Phishing attacks have some common characteristics which can be identified by machine learning methods. In this paper, we compared the results of multiple machine learning methods for predicting phishing websites.

Machine learning technology consists of many algorithms which requires past data to make a decision or prediction on future data. Using this technique, algorithm will analyze various blacklisted and legitimate URLs and their features to accurately detect the phishing websites including zero- hour phishing websites.

Cyber security persons are now looking for trustworthy and steady detection techniques for phishing websites detection. This paper deals with machine learning technology for detection of phishing URLs by extracting and analyzing various features of legitimate and phishing URLs.

v

### TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| **CHAPTER NO.** | **TITLE** | **PAGE NO** |
|  | **ABSTRACT** | **V** |
|  | **LIST OF FIGURES** | **XI** |
| 01 | **INTRODUCTION**   * 1. OVERVIEW OF PROJECT   2. SCOPE OF THE PROJECT | 1  1  1 |
| 02 | **LITERATURE SURVEY**   * 1. INTRODUCTION   2. REVIEW OF LITERATURE SURVEY | 2  2  3 |
| 03 | **SYSTEM ANALYSIS**   * 1. EXISTING SYSTEM      1. Disadvantages   2. PROPOSED SYSTEM      1. Advantages | 7  7  7  8  8 |
| 04 | **REQUIREMENT SPECIFICATIONS**   * 1. INTRODUCTION      1. Functional Requirements      2. Non-Functional Requirements      3. Hardware and Software Requiremen | 9  9  9  9  ts 10 |

4.1.3.1 Software Requirements 10

|  |  |  |
| --- | --- | --- |
|  | 4.1.3.2 Hardware Requirements | 10 |
|  | 4.2 SOFTWARE DESCRIPTION | 10 |
|  | 4.2.1 Anaconda Navigator | 10 |
|  | 4.2.2 Conda | 11 |
|  | 4.2.3 Jupyter Notebook | 12 |
|  | 4.2.3.1 Running the Jupyter Notebook | 12 |
|  | 4.2.3.2 Working Process | 13 |
|  | 4.3 PYTHON | 15 |
|  | 4.3.1 Introduction | 15 |
|  | 4.3.2 Syntax and Semantics | 15 |
|  | 4.3.3 Indentation | 15 |
|  | 4.3.4 Methods | 16 |
|  | 4.3.5 Typing | 16 |
| **05** | **SYSTEM DESIGN** | 17 |
|  | 5.1 ARCHITECTURE DESIGN | 17 |
|  | 5.2 UML DIAGRAMS | 17 |
|  | 5.2.1 Use Case Diagram | 17 |
|  | 5.2.2 Sequence Diagram | 18 |
|  | 5.2.3 Class Diagram | 19 |
|  | 5.2.4 Activity Diagram | 19 |
|  | 5.2.5 Work Flow Diagram | 20 |
|  | 5.2.6 Entity Relationship Diagram (ERD) | 21 |
|  | 5.3 MODULES | 21 |
|  | 5.3.1 Data Pre-Processing | 22 |
|  | 5.3.2 Data Visualization | 23 |

|  |  |  |
| --- | --- | --- |
|  | * + 1. Comparing Algorithm     2. Deployment using Flask | 25  27 |
| **06** | **CODING AND TESTING**   * 1. CODING   2. CODING STANDARDS      1. Naming Conventions      2. Value Conventions      3. Script Writing and Commenting standard   3. TESTING   4. TYPES OF TESTING      1. Unit Testing      2. Functional Testing      3. System Testing      4. Performance Testing      5. Integration Testing      6. Program Testing      7. Validation Testing      8. User Acceptance Testing | 29 |
|  | 29 |
|  | 29 |
|  | 29 |
|  | 29 |
|  | 30 |
|  | 30 |
|  | 30 |
|  | 30 |
|  | 30 |
|  | 31 |
|  | 31 |
|  | 31 |
|  | 32 |
|  | 32 |
|  | 32 |
|  | **CONCLUSION AND FUTURE WORK** | 33 |
|  | **APPENDICES** | 34 |
| APPENDIX A-SOURCE CODE | 34 |
| APPENDIX B-SCREENSHOT | 56 |
| **REFERENCES** | 57 |

## LIST OF FIGURES

|  |  |  |
| --- | --- | --- |
| **FIG NO** | **TITLE** | **PAGE NO** |
| 3.2 | ARCHITECTURE OF PROPOSED MODEL | 8 |
| 5.1 | SYSTEM ARCHITECTURE | 17 |
| 5.2.1 | USECASE DIAGRAM | 18 |
| 5.2.2 | SEQUENCE DIAGRAM | 18 |
| 5.2.3 | CLASS DIAGRAM | 19 |
| 5.2.4 | ACTIVITY DIAGRAM | 19 |
| 5.2.5 | WORK FLOW DIAGRAM | 20 |
| 5.2.6 | ER – DIAGRAM | 21 |

xi