K. S. INSTITUTE OF TECHNOLOGY, BANGALORE-560109

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

PROJECT PHASE 0 + SEMINAR (18CSP77)

2021_CSE_08

" ENHANCING THE PERFORMANCE OF ANTIPHISHING MECHANISM USING MACHINE LEARNING "

Under the guidance of:

Mr. Roopesh Kumar BN

Asst. Professor Dept of CSE

KSIT Bangalore

Presented by:

R Soumya - 1KS18CS125

Sri Chandana P -1KS18CS098

Vijetha - 1KS18CS117

Sushmitha S - 1KS18CS106

CONTENTS

❑ OVERVIEW

❑ GOALS

❑ APPLICATIONS

❑ REQUIREMENT SPECIFICATIONS

❑ DATA SET

❑ REFERENCES

OVERVIEW

- The cyber security problems are increasing nowadays due to the growth of internet world wide . In phishing attack attacker creates a replica of existing link or webpage to fool the user to get access to the personal information.

- Phishers use multiple methods, including email, uniform resource locators (URL),instant messages, forum postings, telephone calls, and text messages to steal user information.

Continued…

- However, detecting phishing is a challenging task, as most of the techniques are not able to make an accurate decision dynamically as to whether the website is phishing or legitimate.

- The ML based phishing techniques depend on website functionalities to gather information that can help classify websites for detecting phishing sites. Here some common supervised learning techniques are applied to accurately detect phishing websites.

GOALS

- Building an application that detects phishing websites efficiently.
- This applications helps us to improve security and provides significant security benefits.
- To get better performance than existing system.
- To prevent the financial loss for victims.
- To prevent identity theft of the online users.

APPLICATIONS

- Increase the user alertness to phishing risk.
- Can be used by e-commerce or other websites.
- Better insights into online behaviour of employees.
- Change behaviour to eliminate the automatic trust response.
- Deploy targeted anti-phishing solutions.
- Protect valuable corporate and personal data.
- Meet industry compliance obligations.

REQUIREMENTS

Software Requirements:

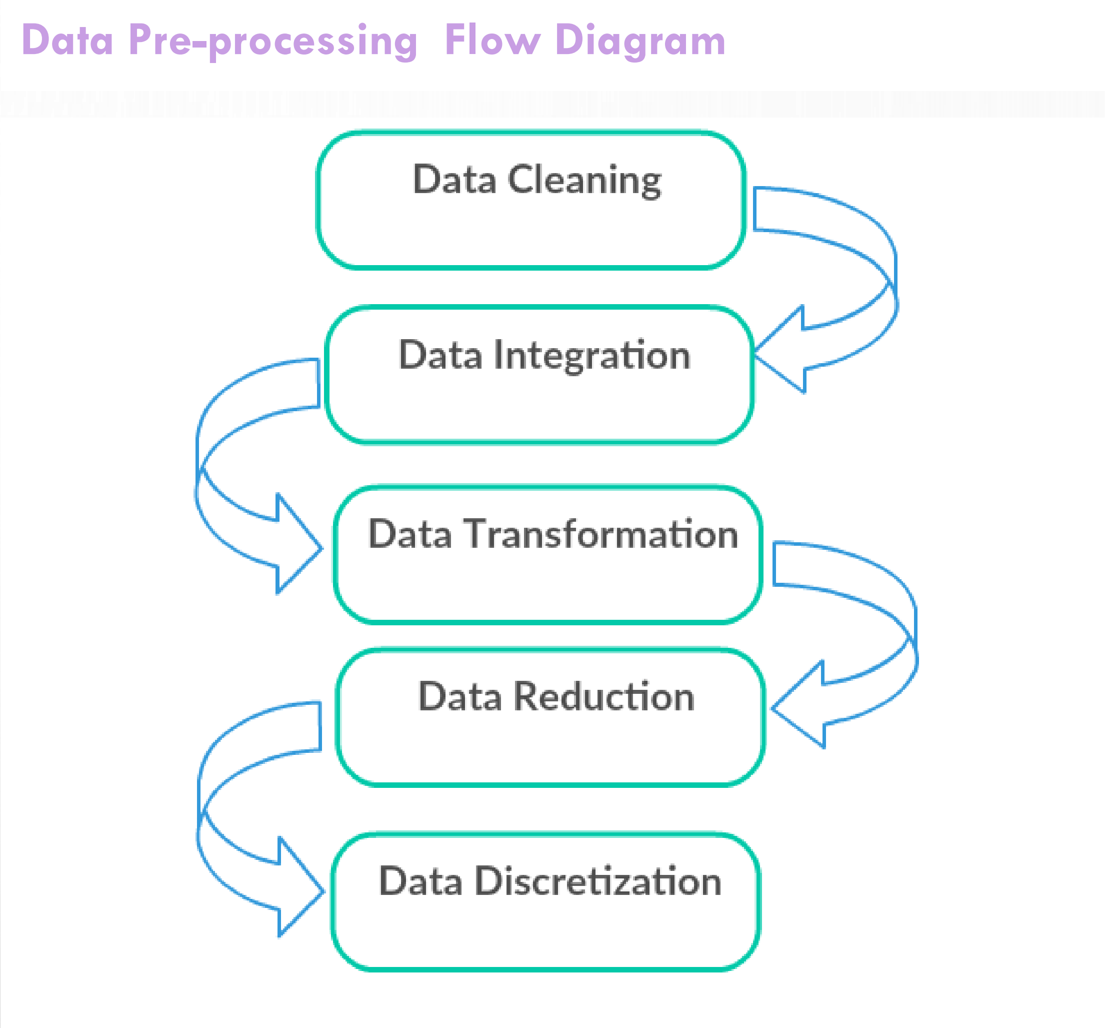
- Jupyter notebook
- Python 3
- JavaScript
- Flask

Hardware Requirements:

- System with i3 processor and above
- 2 GB RAM and above

DATA SET

- ✓ Data collected from the online sources are in the raw form, it may contain errors. It requires correction.
- ✓ Data preprocessing is done to remove outliers and standardize the data.



Data Pre-processing Flow Diagram

Data Cleaning

Data Integration

Data Transformation

Data Reduction

Data Discretization

Continued…

- ✓ Dataset has a records of phishing and legitimate websites.
- ✓ The set of phishing URLs are collected from opensource service i.e, Kaggle and UCI websites.
- ✓ The dataset is divided into 80:20 ratio for training and testing.
- ✓ Dataset contains some category of features that are extracted from the URL of the websites.
- ✓ Set of features denotes website is phished or legitimate.
- ✓ Some extracted features include URL , Page content and Domain based.

REFERENCES

1. Phishing attack detection using feature selection techniques : Aniruddha Narendra Joshi, Tanuja R Pattanshetti, College of Engineering Pune, Wellesley road, Pune, India.
2. Detecting phishing websites using machine learning technique: Ashit Kumar Duttaid, Department of Computer Science and Information System, College of Applied Sciences, Almaarefa University, Riyadh, Saudi Arabia.

THANK YOU