PHISHING DOMAIN DETECTION

(Machine Learning)

HIGH LEVEL DESIGN (HLD)

INTRODUCTION:

Phishing Domain Detection is a technique by which we should be able to predict whether the domain is real or fake.

PROBLEM STATEMENT:

Phishing is a type of fraud in which an attacker impersonates a reputable company or person in order to get sensitive information such as login credentials or account information via email or other communication channels. Phishing is popular among attackers because it is easier to persuade someone to click a malicious link that appears to be authentic than it is to break through a computer's protection measures.

The mail goal is to predict whether the domains are real or malicious.

APPROACH:

The classical machine learning tasks like Data Exploration, Data Cleaning,

Feature Engineering, Model Building and Model Testing as been done on the project. Tried with different machine learning algorithms such as Logistic Regression, SVM, Gradient Boosting, KNN, Random Forest and found out best fit model in the project as Random Forest.

DATASET:

These data consist of a collection of legitimate as well as phishing website instances. Each website is represented by the set of features which denote, whether website is legitimate or not. Data can serve as an input for machine learning process.

The dataset had two variants of the Phishing Dataset are presented.

Full variant - dataset_full.csv

- Short description of the full variant dataset:
- Total number of instances: 88,647
- Number of legitimate website instances (labeled as 0): 58,000
- Number of phishing website instances (labeled as 1): 30,647
- Total number of features: 111

Small variant - dataset_small.csv

- Short description of the small variant dataset:
- Total number of instances: 58,645
- Number of legitimate website instances (labeled as 0): 27,998
- Number of phishing website instances (labeled as 1): 30,647
- Total number of features: 111

TOOLS USED:

Python Programming languages and libraries such as NumPy, Pandas, Matplotlib, Seaborn, Scikit learn were used to build the whole model and Flask were used for a web framework.

1. Data collection and preparation

The first step is to collect a dataset of phishing and legitimate domains. This dataset can be gathered from a variety of sources, such as public phishing datasets, honeypots, and web scraping. Once the dataset is collected, it needs to be prepared for training the machine learning model. This includes cleaning the data, removing duplicates, and extracting features.

2. Feature extraction

The next step is to extract features from the dataset. Features are the characteristics of a domain that can be used to distinguish between phishing and legitimate domains. Some common features include the length of the domain name, the presence of hyphens or underscores, the use of top-level domains (TLDs) that are commonly associated with phishing, and the number of typos in the domain name.

3. Machine learning model training

Once the features are extracted, the machine learning model can be trained. The model is trained on a dataset of phishing and legitimate domains, and it learns to predict whether a new domain is phishing or legitimate. There are a variety of machine learning algorithms that can be used for phishing domain detection, such as decision trees, random forests, and support vector machines.

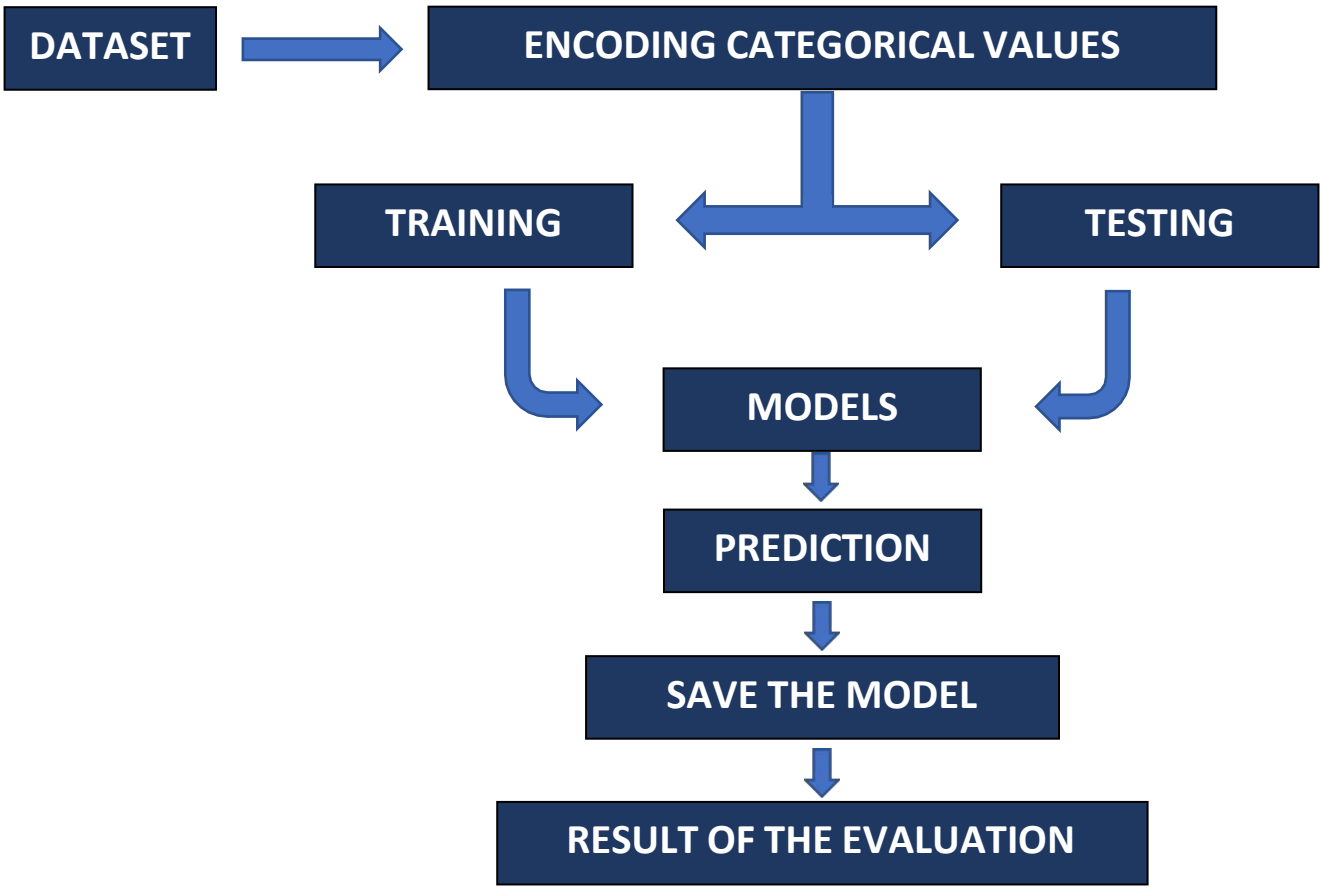
4. Model evaluation and deployment

Once the model is trained, it needs to be evaluated. This is done by testing the model on a held-out dataset of phishing and legitimate domains. The evaluation results will help to determine the accuracy of the model. Once the model is evaluated, it can be deployed to production.

5. Ongoing maintenance

The phishing domain detection system needs to be maintained on an ongoing basis. This includes updating the dataset with new phishing domains, retraining the model, and monitoring the performance of the mode

DESIGN FLOW:


DATASET
ENCODING CATEGORICAL VALUES
TRAINING
TESTING
MODELS
PREDICTION
SAVE THE MODEL
RESULT OF THE EVALUATION

CONCLUSION:

In conclusion, the results obtained from the phishing domain detection using the random forest model have been quite promising. With an accuracy of 98%, precision of 98%, and recall of 97.6%, the model has demonstrated excellent performance in correctly identifying phishing domains. These metrics indicate that the model has a high degree of accuracy in detecting malicious domains and is reliable in predicting whether a domain is phishing or not. Overall, this model can be a valuable tool in protecting users against phishing attacks, and its high accuracy and precision make it a strong candidate for use in real-world scenarios.