

# Basic Details of the Team and Problem Statement




Problem Statement ID: 14

Problem Statement title/Domain :

**Create an intelligent system using AI/ML to detect phishing domains which imitate look and feel of genuine domains.**



# Describe your idea/Solution/Prototype here:



Phishing is the most common Cyber Attack often used to steal the data of the user. Fake links/websites are shared through number of mediums like email, SMS etc. To target users. These domains host user login page that imitates the genuine target websites.

Login attempts on such pages can lead to compromise of user credentials and may also download malicious payload in user computers.

## **OUR SOLUTION:**

To tackle this problem we aim to use machine learning, Our solution involves collecting datasets of phishing and legitimate URLs, extracting key features from these URLs (such as length, number of hyphens, page rank, web Traffic, domain information etc), and train machine learning model to differentiate between phishing and legitimate URLs.

The model is then integrated into a web application. This web app allows users to input URLs and receive real-time predictions on their legitimacy and gives the necessary description about it.

# Idea/Approach Details

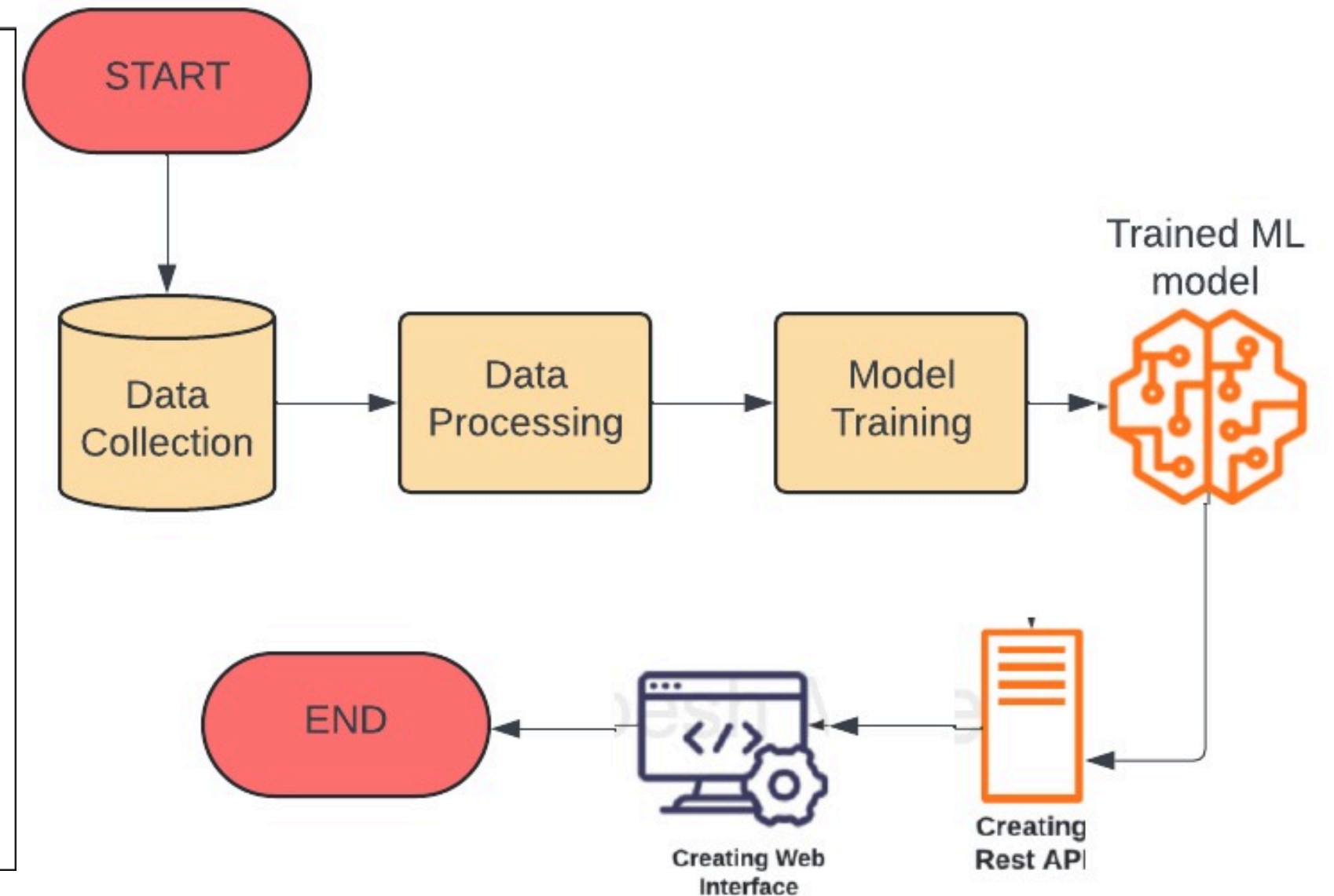
Describe your Technology stack here:

**Programming languages:** Python

**Machine Learning Algorithm:** Random forest Classifier.

**Libraries and frameworks:**

Scikit-learn, TensorFlow, PyTorch, Pandas, Numpy,  
For webapplication: Flask, HTML/CSS/JavaScript



# Idea/Approach Details

- End Users : General public ,cyber Security teams (Fake links/websites are shared through number of mediums like email, SMS etc)

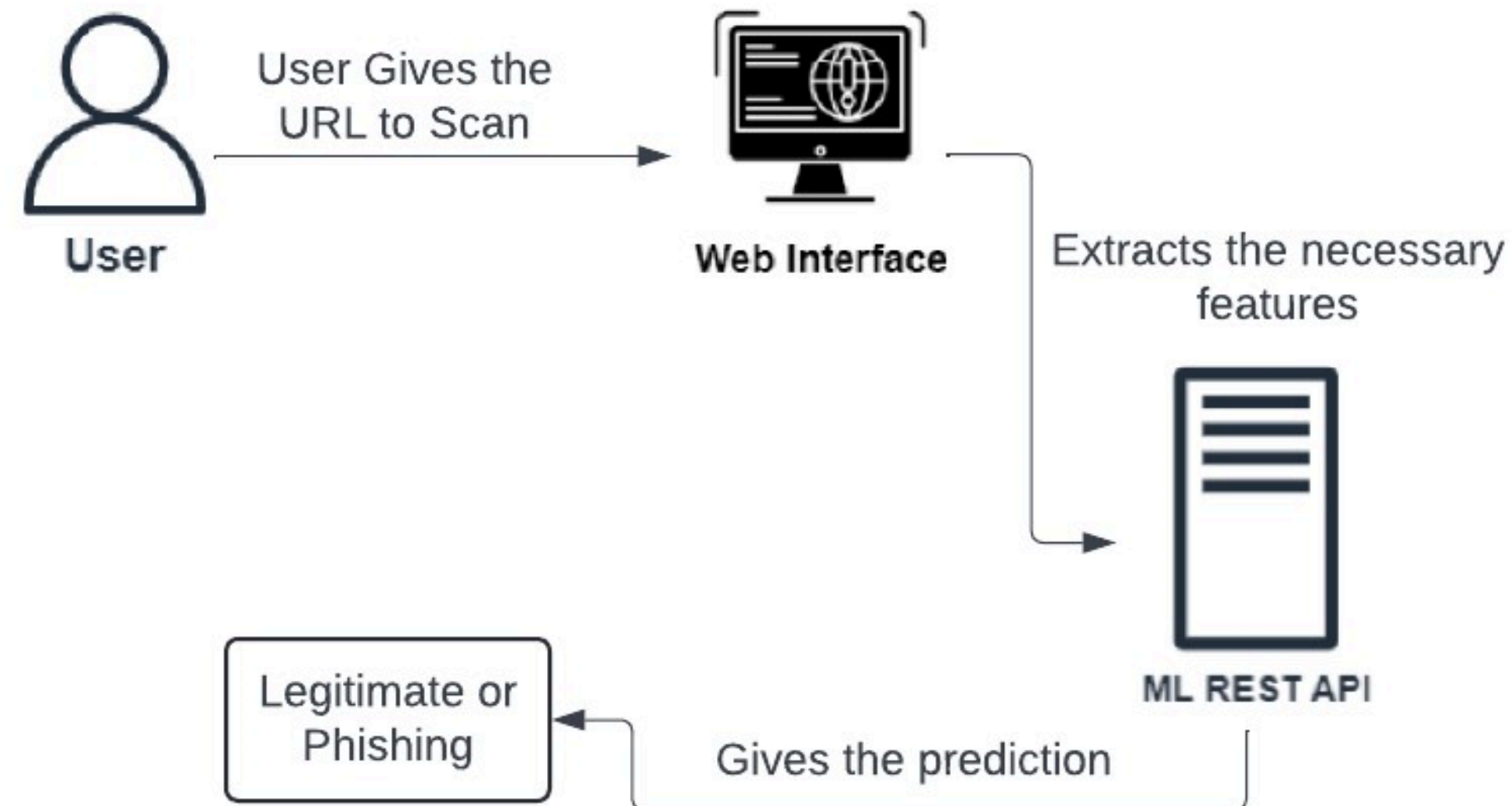
## 1. User Interaction:

1. A user accesses the web application via their web browser.
2. The user inputs a URL into the provided form.
3. The user submits the URL for analysis.

## 1. Show Stopper:

- Access to a large, high-quality, labeled dataset of phishing and legitimate URLs is crucial.
- Incorrect or incomplete feature extraction can lead to poor model performance.
- Regular updates to the model and application are required to keep up with evolving phishing techniques.

# Prototype projection and conclusion



**Conclusion:** Our approach leverages machine learning and modern web technologies to provide a scalable and user-friendly solution for enhancing cybersecurity and protecting users from phishing attacks.

[Home](#)[Predict](#)[Report URL](#)[Login](#)

# CatchPhish

Don't get hooked by  
phishers.



# CatchPhish

Enter URL

Scan

## CatchPhish

**Domain is Safe**

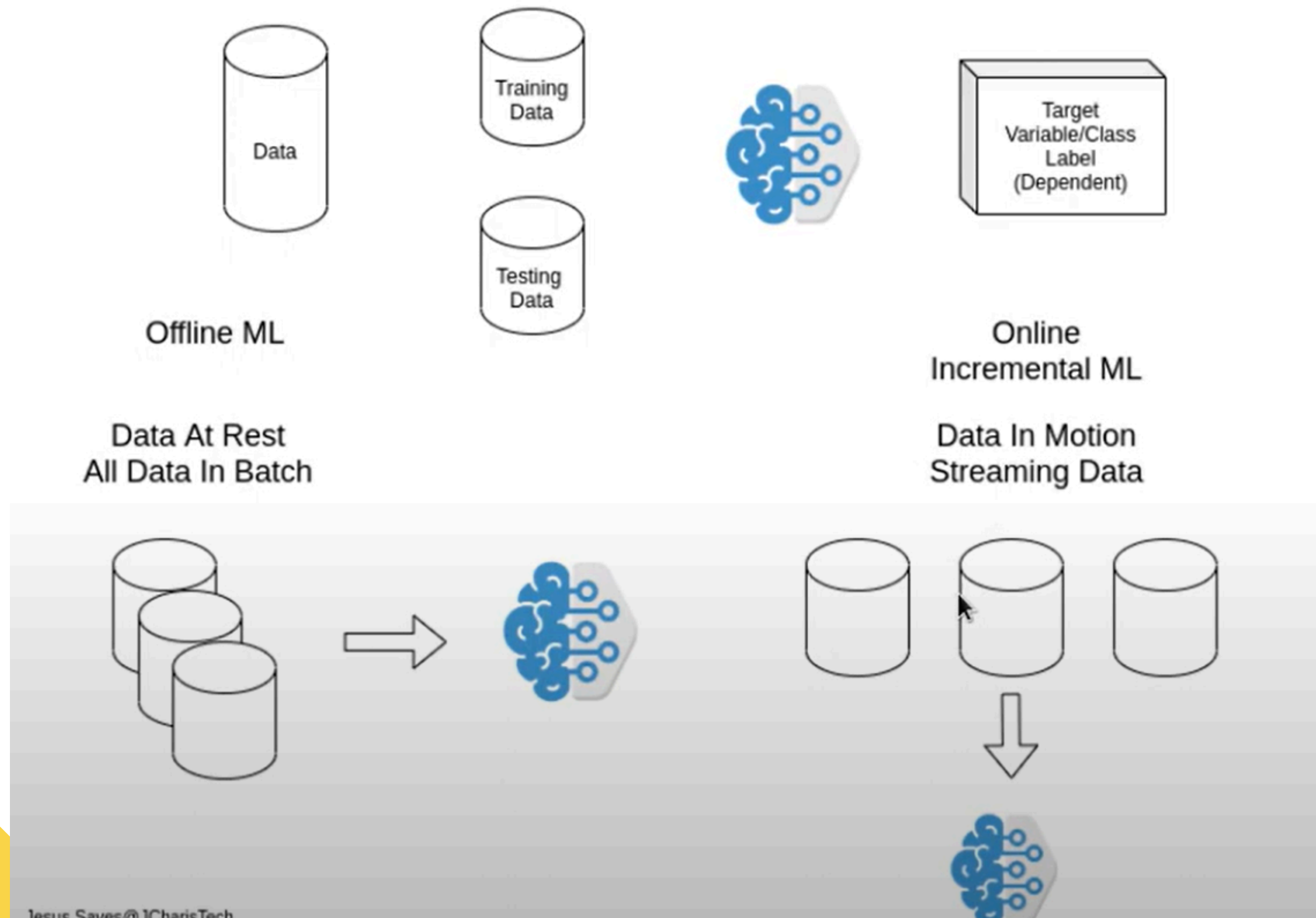
This URL is 85% likely to be legitimate.

## CatchPhish

**Domain is Unsafe**

This URL is 90% likely to be phishing.

# Future Scope: Incremental Machine Learning







THANK YOU