

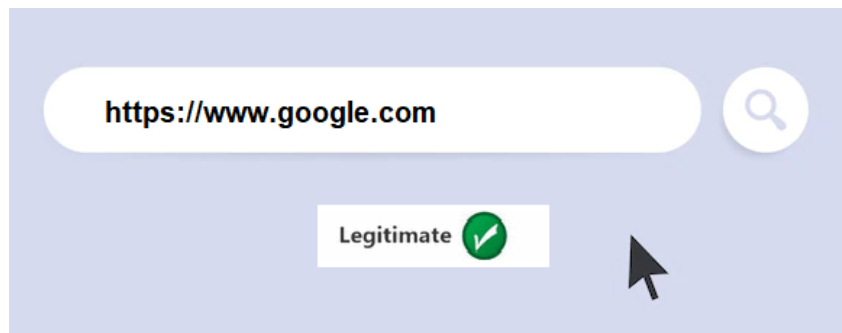
Project Design Phase-I
Solution Architecture

Date	19 September 2022
Team ID	PNT2022TMID15077
Project Name	Web Phishing Detection
Maximum Marks	4 Marks

Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- We can use various machine learning algorithms to detect web phishing.
- At first, we need to preprocess the dataset provided and also detect the important features using feature importance and also by using feature selection techniques which can be done by inbuilt libraries.
- We are able to split the dataset into training and testing.
- We train the dataset on several algorithms such as logistic regression, naive bayes, decision tree, support vector machine etc.,
- We need to test the models with the test dataset and calculate the accuracy of those models.
- So, With higher accuracy model we are able to use the model as the default model and predict the URL to be safe or not.
- UI has been created by web developers, where the user can give a URL in the website in the text box provided in the .
- The model will be stored in the cloud. The website will get the required parameters from the URL and pass it to the model prediction using API. This will predict the URL is legitimate or phishing.



- Those results will be sent to the user on the website.

Solution Architecture Diagram:

