**Phishing Website Detection**

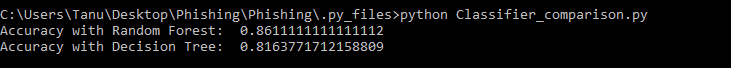
**Aim of the project :**

To detect the phishing websites with high accuracy.

**To do so :**

First, we have chose the best algorithm suited for project :

1. Decision tree
2. Random forest

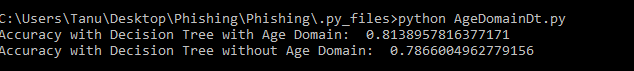


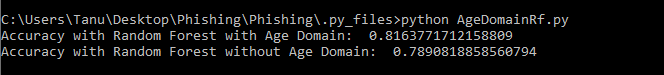
Second, we have selected the features to increase the accuracy of the model :

1. **Age of Domain** – This feature can be extracted from WHOIS directory . Most phishing websites live for short period of time.

Age of domain >= 6 months 🡪 legitimate

Otherwise 🡪 phishing



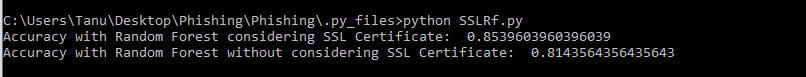


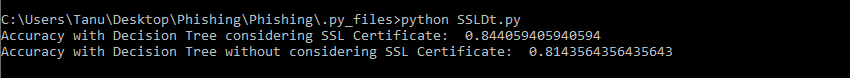
1. **HTTPS (Hypertext transfer protocol with secure socket) :**

Use https and issuer is trusted and age of certificate >= 1years 🡪 legitimate

Using https and issueris not trusted 🡪 suspicious

Otherwise🡪 phishing





Overall accuracy improved by adding the above 2 features together:



