MALICIOUS WEBSITE NAVIGATION PREVENTION USING CNNs AND URL VECTORS: A STUDY

Guide Name	Panel Head
Dr K. Kalaiselvi	Dr K. Kalaiselvi

Faculty Advisor	Project Domain
Dr M. Uma	Cyber Security

Student(s) Details: Name

1. Syed Abbas Haider Rizvi
2. Siddhant Tiwari

Passport size photo(s)

Registration Number(s)

1. RA1811030010066
2. RA1811030010082

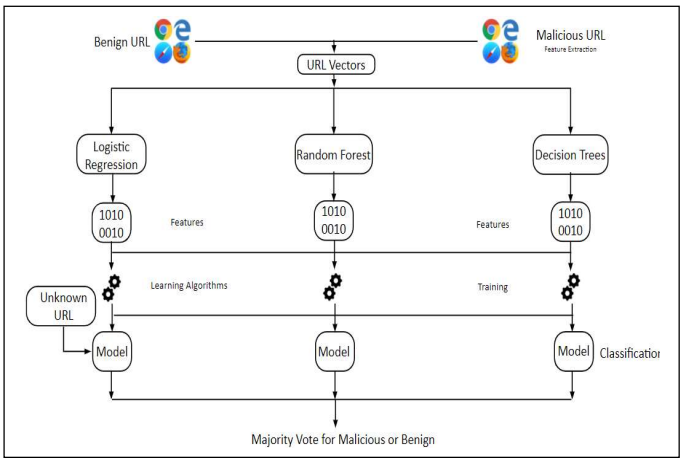
Email ID(s)&Mobile Number(s)

1: st6535@srmist.edu.in	2: sr5457@srmist.edu.in
8237807125	9595032334

Abstract

In this project, we have focused on the problem of malicious URLs. URL attacks have been on the rise in 2020, with most of the work being online based due to the pandemic there arises a greater scope of Phishing URLs etc. There have been existing systems, but they are mostly paid, whereas with this project we aim to deploy a free website/ URL checker with a good amount of accuracy and to compare the accuracies for CNNs and other available models.

Architecture Diagram



Significance of the Project

1)Case study on ML models and other efficient methods to resolve the problem of URL phishing.
2)Testing the ML models with variety of datasets.
3)Comparing URL vectors and features against each other and discovering the important ones.

Conclusion

As it is evident from the classification report that the CNN model provides us with great accuracy when the dataset concerns non https/http URLs. The accuracy achieved here using the URL vectors is quite better when considering the dataset involving http/https URLs
So we can conclude that because of a better confusion matrix CNN model may prevail but the time/space factors are concerning.

Conference/Journal Publication Details (If Any)

ICCCI (Paper Submitted Jan 2022) - Paper Published April 2022

ICBCC (Paper Submitted May 2022)