WEB PHISHING DETECTION

USING APPLIED DATA SCIENCE

TEAM ID : PNT2022TMID44694

TEAM MEMBERS : DHIVYA S (732519104004)

ABITHA J (732519104002)

VASUNTHARA R (732519104030)

PRIYATHARSHINI G (732519104703)

DEPARTMENT : CSE

COLLEGE NAME : SHREE VENKATESHWARA HI-TECH ENGINEERING COLLEGE, GOBI.

PROPOSED SOLUTION FIT

The overview of the proposed solution to detect phishing attacks is the data source contained URLs and HTML codes of web pages. The URLs are directly used as inputs to the model with a minimum pre-processing, and that is separately discussed in a below subsection. However, HTML features need to be extracted from the web pages. Therefore, a feature extraction model is used for the extraction before finalizing the model input features. After extracting the relevant features from the web pages, HTML features, and URLs concatenate to have input feature vectors for the detection model. Finally, the detection model will use the input feature vector and produce an output as legitimate or phishing. However, the

detection model is a combination of two deep networks. It can analyze URLs and HTML features separately and combine both decisions in making the final output of the model. The major components included in the solution, namely, a feature extraction model and detection model, are introduced in the following subsections.

PROPOSED SOLUTION TO DETECT PHISHING ATTACKS