# IDEATION

The process of dataset processing, feature selection, and dataset division was presented in Chapter 4. This chapter addresses the problem of selecting the best classification technique for website phishing detection that causes degradation in detection accuracy and high false alarm rate.

Mustafa Aydin et al. proposed a classification algorithm for phishing website detection by extracting websites' URL features and analyzing subset based feature selection methods. It implements feature extraction and selection methods for the detection of phishing websites. This is an interactive and responsive website that will be used to detect whether a website is legitimate or phishing. This website is made using different web designing languages which include HTML, CSS, Javascript and Django.

One of the major contributing factors to low overall accuracy is the selection of weak weighted features for classification. The situation worsens when a lazy algorithm is trained and tested with a large dataset. Therefore, the performance of the research methodology used in this project may not perform so well if the wrong classifier is trained and tested with dataset size more than the classifier’s capacity. Nowadays Phishing becomes a main area of concern for security researchers because it is not difficult to create the fake website which looks so close to legitimate website. Experts can identify fake websites but not all the users can identify the fake website and such users become the victim of phishing attack. Main aim of the attacker is to steal banks account credentials. Phishing attacks are becoming successful because lack of user awareness. Since phishing attack exploits the weaknesses found in users, it is very difficult to mitigate them but it is very important to enhance phishing detection techniques.