Phishing attacks have become a significant threat to online security, as cybercriminals try to trick unsuspecting users into revealing sensitive information through fraudulent emails, websites, and messages. In this project, we proposed a machine learning-based approach for detecting phishing attacks using the random forest classification.

Our Project's main ambition is to bring AI powered phishing detection solution to every layman irrespective of their device specifications. This solution is divided into two parts - backend and the frontend that co-exist with each other.

The backend will remain in a web server. Here Arff datasets will be preprocessed and features and attributes of the dataset are extracted into a numpy array.This array is then fed into the random forest classifier which in turn creates the trained model.This trained model in the form of a JSON fle is hosted on the server for users to download into their devices.In future iterations we will create a reporting systems in which users can report a phishing site (after which that site's data will be fed into our model).

The frontend is a web plugin installed on the user's device . On the first run it will download the cached model from our web server. Then everytime a site opens, it will extract the site's features and then use the model to figure out if it is a phishing site or not. Appropriately it will show a popup on the user's browser.

We are using a dataset created by UCI which has 30 different attributes and over 2456 instances of data.

After preprocessing and feature extraction, we trained a random forest classifier on our dataset using 10-fold cross-validation. The model achieved an accuracy of over 94.7401% on the test set, outperforming several other classifiers such as logistic regression and support vector machines. We also evaluated the model's performance on a real-world phishing dataset and found that it could detect over 90% of the attacks with a low false positive rate.

To further validate the effectiveness of our approach, we conducted a comparative analysis with other state-of-the-art phishing detection techniques, such as rule-based systems and directory based systems. Our results showed that the random forest algorithm was able to achieve higher accuracy and faster execution times than these methods, making it a promising approach for real-time phishing detection.

Overall, our project demonstrated the potential of machine learning in improving online security by automating the detection of phishing attacks. The random forest algorithm, in particular, showed great promise in accurately identifying fraudulent communications and reducing the risk of users falling victim to such attacks. Our work can serve as a basis for further research in this area and contribute to the development of more effective cybersecurity measures.