**NLP-Based Phishing Email Detection**

**ABSTRACT**

Managing email interactions is becoming increasingly difficult for both individuals and organizations due to the rise of unwanted and destructive emails, or spam. As a result, detecting spam emails has become essential for preserving email security, increasing productivity, and lowering the danger of online risks including malware, phishing, and identity theft. Using machine learning techniques, this project seeks to create an efficient spam email detection system that can distinguish between real and spam emails.

The system makes use of a number of characteristics that are taken from the email content, such as the metadata, sender details, body text, and subject. A number of machine learning techniques, such as Random Forest, Support Vector Machines (SVM), and Naive Bayes, are used and evaluated on how well they classify emails.

A large range of emails, both spam and valid, are included in the dataset used for training and assessment. In order to enhance classification performance, the project also investigates methods such as feature extraction, text preparation, and natural language processing (NLP). The ultimate objective of this project is to create an effective spam filter that can lower the volume of unsolicited and potentially dangerous emails, guaranteeing a safer, cleaner, and more effective email communication experience.