**Spam email detection using NLP**

Abstract:

Spam email detection using Natural Language Processing (NLP) is a project aimed at classifying emails into spam or non-spam categories to enhance communication security and efficiency. This system employs NLP techniques to preprocess and analyze text data from email content, extracting meaningful features such as keywords, tone, and grammar. Machine learning models like Logistic Regression, Naïve Bayes, and Deep Neural Networks are trained on labeled datasets to classify emails accurately.

Advanced methods such as n-grams, word stemming, and feature engineering are utilized to capture contextual patterns in email text. Additionally, deep learning models like LSTM and BiLSTM improve detection by understanding sequential dependencies in text. Visualization tools like confusion matrices further evaluate model performance. This project addresses challenges like subtle differences between spam and legitimate emails, offering a robust solution for automated email filtering.