20MCA245 Mini Project Synopsis
Approval by Guide

Name of Student: Shijna M

Roll No.: 51

Project Title: Spam Mail Prediction Using Machine Learning

Abstract: Email communication plays a fundamental role in modern society, but the influx of unwanted or spam emails can hinder its efficiency. Logistic regression model for the prediction and classification of incoming emails as spam or ham. Proposed model utilizes a set of carefully chosen features, including word frequency analysis, sender information, subject line characteristics, and the presence of attachments. These features are extracted from the email content and headers, providing a comprehensive view of the email's attributes. A labeled dataset, consisting of both spam and ham emails, is employed to train the logistic regression model. Evaluation metrics such as accuracy, precision, recall, and F1-score are employed to assess the model's performance. Experiments conducted on email dataset demonstrate the model's effectiveness in accurately predicting and filtering spam emails, thereby improving email communication efficiency and user experience.

Keywords: Spam, Logistic Regression model, Evaluation metrices.

Signature of Student:
Any remarks of Guide:

Name and Signature of Guide: Prof. Ancy Emmanuel
Date: 04/10/2023