Spam Mail Detection

Project Overview:

The primary objective of this project is to classify whether the email is spam or not. The dataset comprises one feature and one target variable, containing 5,727 rows and 2 columns.

Impact:

Through extensive exploratory data analysis, we have unearthed significant insights. The email text lengths range from a minimum of 13 to a maximum of 43,952 . A comparison between spam and non-spam messages revealed 1,368 instances of spam messages and 4,360 non-spam messages. Additionally, unique characteristics within each message were identified.

For classification modeling, employed the Naive Bayes Classifier, which yielded an impressive accuracy rate of 99%. This project provides valuable insights to email users, empowering them to make informed decisions based on email types.