****

**Assessment Report**

on

**“Spam Emails Detection”**

submitted as partial fulfillment for the award of

**BACHELOR OF TECHNOLOGY**

**DEGREE**

SESSION 2024-25

in

**CSE(AI)**

By

Name : Himalaya Vashistha

Roll Number : 202401100300121

Section: B

**Under the supervision of**

“Shivansh Prasad”

**KIET Group of Institutions, Ghaziabad**

**May, 2025**

**1. Introduction**

As email remains a primary mode of communication, the rise of spam poses serious challenges to cybersecurity and user experience. Automating the detection of spam emails through data-driven approaches is essential in combating unwanted or malicious messages. This project focuses on identifying spam emails using supervised machine learning techniques. By leveraging a dataset containing features such as the number of links, attachments, and sender reputation, the objective is to build a predictive model that accurately classifies emails as spam or not spam, thereby enhancing email security and efficiency.

**2. Problem Statement**

To predict whether an email is spam or not using features such as the number of links, attachments, and sender reputation. This classification will help email systems reduce spam exposure and enhance user security by accurately identifying and filtering out unwanted or harmful messages.

**3. Objectives**

● Preprocess the email dataset for training a machine learning model.

● Train a Random Forest classifier to identify spam emails.

● Evaluate model performance using standard classification metrics such as accuracy, precision, recall, and F1-score.

● Visualize the confusion matrix using a heatmap for improved interpretability.

**4. Methodology**

● **Data Collection**:  
The user uploads a CSV file containing email data with features such as the number of links, attachments, and sender reputation.

● **Data Preprocessing**:

○ Mapping categorical labels ("yes"/"no") in the target variable to binary values (1/0).  
○ Checking and handling any missing values if present.  
○ Ensuring all features are in a suitable numerical format for model training.

● **Model Building**:

○ Splitting the dataset into training and testing subsets using an 80-20 split.  
○ Training a Random Forest classifier on the training data.

● **Model Evaluation**:

○ Evaluating the model using classification metrics including accuracy, precision, recall, and F1-score.  
○ Generating a confusion matrix and visualizing it using a heatmap for better interpretability.

**5. Data Preprocessing**

The dataset is cleaned and prepared as follows:

●The target variable is\_spam is converted from categorical values ("yes"/"no") to binary values (1 for spam, 0 for not spam).

● The dataset is checked for missing values; if present, numerical columns are filled using the mean.

● As all features are numerical, one-hot encoding is not required.

● Feature values are scaled using StandardScaler to ensure uniformity across input variables.

● The dataset is split into 80% for training and 20% for testing.

**6. Model Implementation**

Random Forest is used due to its robustness and high performance in classification tasks. The model is trained on the processed dataset and used to predict whether an email is spam or not on the test set.

**7. Evaluation Metrics**

The following metrics are used to evaluate the model:

● **Accuracy**: Measures overall correctness of the predictions.  
● **Precision**: Indicates the proportion of predicted spam emails that were actually spam.  
● **Recall**: Shows the proportion of actual spam emails that were correctly identified.  
● **F1 Score**: Harmonic mean of precision and recall, giving a balance between the two.  
● **Confusion Matrix**: Visualized using a Seaborn heatmap to analyze prediction errors and performance.

**8. Results and Analysis**

● The model demonstrated strong performance in classifying spam versus non-spam emails.  
● The confusion matrix heatmap helped visualize the distribution of true positives, true negatives, false positives, and false negatives.  
● Precision and recall highlighted the model's effectiveness in detecting spam emails while minimizing false positives.

**9. Conclusion**

The Random Forest model successfully classified spam emails with promising performance metrics. This project illustrates the power of machine learning in enhancing email filtering systems and strengthening digital communication security. Future improvements could include testing other algorithms, tuning hyperparameters, and addressing potential class imbalances.

**10. References**

● scikit-learn documentation  
● pandas documentation  
● Seaborn visualization library  
● Research articles on spam detection and email classification

