|  |  |  |
| --- | --- | --- |
| Name | Roll No | Class |
| Siddhant Sanjay Sathe | 50 | D15A |
| Arnav Santosh Sawant | 52 | D15A |
| Pranav Pramod Titambe | 60 | D15A |

**Topic:- Email Spam Classifier: A Machine Learning Approach**

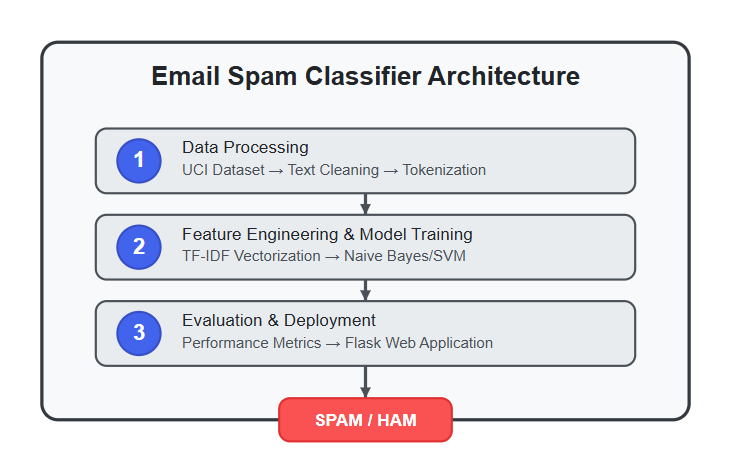
**Introduction:-**   
Email spam, also known as junk email, has become a pervasive issue, causing significant inconvenience and security risks. Spam emails often contain phishing attempts, malware, or fraudulent schemes, posing a threat to individuals and organizations. Machine learning techniques have emerged as a powerful tool for automatically detecting and filtering spam emails. This report explores the use of machine learning in building email spam classifiers. It includes a review of relevant literature, a discussion of the methodology involved, a system diagram, and an overview of the advantages, disadvantages, applications, and objectives of such systems.

**Research papers:-**

|  |  |  |  |
| --- | --- | --- | --- |
| **Paper Name** | **Date** | **Abstract** | **Key Points** |
| **"A comparative study of machine learning methods for email spam detection"** | Dec, 2020 | This paper compares several machine learning algorithms, including Naive Bayes, Support Vector Machines (SVM), and Random Forests, for email spam detection. The performance of each algorithm is evaluated based on metrics such as accuracy, precision, and recall. | - Comparative analysis of multiple ML algorithms. - Performance evaluation using standard metrics. - Highlights the strengths and weaknesses of each algorithm for spam detection. |
| **"Spam Email Detection based on Machine Learning Techniques"** | May, 2023 | This study proposes a spam email detection system using machine learning techniques such as Naive Bayes and Support Vector Machine (SVM). The system is trained on a dataset of labeled emails and uses feature extraction techniques to identify relevant characteristics of spam emails. The results demonstrate the effectiveness of the proposed system. | - Feature extraction techniques for identifying spam characteristics. - Training and evaluation of ML models on labeled email datasets. - Focus on Naive Bayes and SVM algorithms. |
| **"An efficient email spam detection system using hybrid feature selection and machine learning techniques"** | June, 2024 | This paper introduces an efficient email spam detection system that combines hybrid feature selection and machine learning techniques. The proposed system uses a combination of filter and wrapper-based feature selection methods to select the most relevant features for spam detection. | - Hybrid feature selection methods for improved accuracy. - Focus on efficiency and reducing computational complexity. - Integration of feature selection with machine learning algorithms. |

**Methodology:-**   
The methodology for developing an email spam classifier typically involves the following steps:

1. **Data Collection**: Gather a dataset of emails, labeled as either spam or ham (non-spam). Publicly available datasets like the SpamAssassin dataset or the Enron-Spam dataset can be used.
2. **Data Preprocessing**: Clean and preprocess the email data. This includes:
   * Removing HTML tags and special characters.
   * Converting text to lowercase.
   * Removing stop words (e.g., "the," "a," "is").
   * Stemming or lemmatization to reduce words to their root form.
3. **Feature Extraction**: Extract relevant features from the preprocessed email text. Common features include:
   * **Bag of Words (BoW)**: Representing the email as a collection of individual words and their frequencies.
   * **Term Frequency-Inverse Document Frequency (TF-IDF)**: Weighing words based on their frequency in the email and their rarity across the entire dataset.
   * **N-grams**: Considering sequences of n words to capture contextual information.
   * **Header Features**: Analyzing email headers for suspicious patterns (e.g., forged sender addresses).
4. **Model Training**: Train a machine learning model using the extracted features and labeled data. Popular algorithms include:
   * **Naive Bayes**: A probabilistic classifier based on Bayes' theorem.
   * **Support Vector Machines (SVM)**: A powerful classifier that finds the optimal hyperplane to separate spam and ham emails.
5. **Model Evaluation**: Evaluate the performance of the trained model using metrics such as:
   * **Accuracy**: The overall percentage of correctly classified emails.
   * **Precision**: The percentage of emails correctly identified as spam out of all emails predicted as spam.
   * **Recall**: The percentage of emails correctly identified as spam out of all actual spam emails.
   * **F1-score**: The harmonic mean of precision and recall.
6. **Deployment**: Deploy the trained model to filter incoming emails in real-time.

**Diagram:-   
**

**Advantages:-**

* **High Accuracy**: Machine learning models can achieve high accuracy in detecting spam emails by learning complex patterns and relationships in the data.
* **Automation**: Automated spam filtering saves users time and effort by automatically classifying and filtering unwanted emails.
* **Adaptability**: Machine learning models can adapt to new spam techniques and trends by continuously learning from new data.
* **Personalization**: Models can be personalized to individual user preferences and email usage patterns, improving filtering accuracy.

**Disadvantages:-**

* **Data Dependency:** The performance of machine learning models heavily depends on the quality and quantity of training data.
* **Bias:** Models can inherit biases from the training data, leading to unfair or inaccurate filtering.
* **Computational Cost:** Training and deploying complex machine learning models can be computationally expensive.
* **False Positives:** There is a risk of misclassifying legitimate emails as spam (false positives), which can be frustrating for users.
* **Evasion Techniques:** Spammers may develop techniques to evade detection by machine learning models, requiring continuous model updates and improvements.

**Applications**

* **Email Service Providers:** Use spam filters to protect users from unwanted and malicious emails.
* **Corporate Email Systems:** Implement spam filters to safeguard sensitive information and prevent phishing attacks.
* **Individual Email Clients:** Offer built-in or add-on spam filters to help users manage their inbox.
* **Security Software:** Integrate spam filtering capabilities to provide comprehensive protection against online threats.

**Objectives**

* To collect and preprocess a diverse dataset of email messages labeled as spam or ham.
* To extract relevant features from email content and headers using techniques like TF-IDF, and N-grams.
* To train and evaluate machine learning models, such as Naive Bayes, SVM, and Random Forest, for email spam classification.
* To optimize model parameters and feature selection methods to achieve high accuracy, precision, and recall.
* To develop a system that can effectively filter spam emails while minimizing false positives.