SPAM MAIL DETECTION USING NAIVE BAYES USING MACHINE LEARNING

K.VIMAL PANDIYAN
Nehru Institute Of Engineering And Technology
Department of CSE

P.JEYA SANTHOSH
Nehru Institute Of Engineering And Technology
Department Of CSE

R.SHATHIYAPRIYAN
Nehru Institute Of Engineering And Technology
Department Of CSE

Abstract

Nowadays, emails are used in almost every field, from business to education. Emails have two subcategories, i.e., ham and spam. Email spam, also called junk emails or unwanted emails, is a type of email that can be used to harm any user by wasting his/her time, computing resources, and stealing valuable information. The ratio of spam emails is increasing rapidly day by day. Spam detection and filtration are significant and enormous problems for email and IoT service providers nowadays. Among all the techniques developed for detecting and preventing spam, filtering email is one of the most essential and prominent approaches. Several machine learning and deep learning techniques have been used for this purpose, i.e., Naïve Bayes, decision trees, neural networks, and random forest. This paper surveys the machine learning techniques used for spam filtering techniques used in email and IoT platforms by classifying them into suitable categories. A comprehensive comparison of these techniques is also made based on accuracy, precision, recall, etc. In the end, comprehensive insights and future research directions are also discussed. Email spam detection is an important task in the field of natural language processing, as it helps in identifying unwanted emails and prevents them from reaching the user's inbox. One of the popular approaches for email spam detection is the Naive Bayes algorithm. In this paper, we propose a Naive Bayes-based approach for email spam detection. The proposed approach involves preprocessing the email data, feature extraction, and then applying the Naive Bayes algorithm for classification. We use a publicly available dataset to evaluate the performance of the proposed approach. The results show that our approach achieves a high accuracy rate in detecting spam emails. Therefore, the proposed approach can be used as an effective method for email spam detection, which can help in improving the overall user experience and reducing the risk of security breaches.

Keywords—Machine Learning, Spam Mail Detection, Naive Bayes Algorithm, Accurate Classification, Predictive Modeling

1. INTRODUCTION

Email is a popular medium for communication, but it also poses a significant risk to users' security and privacy due to the high volume of spam emails. Spam emails are unwanted emails that are sent to many recipients for advertising or other malicious purposes. Email spam can lead to several problems, including phishing attacks, malware distribution, and identity theft. Therefore, email spam detection is crucial in preventing these risks. Machine learning algorithms, especially Naive Bayes, have been widely used in email spam detection. Naive Bayes is a probabilistic algorithm that uses Bayes' theorem to calculate the probability of a given email being spam or not. The Naive Bayes algorithm is simple, fast, and effective for text classification tasks, making it a popular choice for email spam detection. In this paper, we propose a Naive Bayes-based approach for email spam detection. The proposed approach involves preprocessing the email data, feature extraction, and then applying the Naive Bayes algorithm for classification. We use a publicly available dataset to evaluate the performance of the proposed approach. The results show that our approach achieves a high accuracy rate in detecting spam emails. Therefore, the proposed approach can be used as an effective method for email spam detection, which can help in improving the overall user experience and reducing the risk of security breaches. I hope this email finds you well. I am writing to provide you with an introduction to the concept of Naive Bayes and its application in email spam detection.

1.1 BACKGROUND

Email spam detection has been an active research area in the field of natural language processing for several years. With the increasing volume of emails being sent every day, the need for an effective method to detect and filter out spam emails has become more critical. Several approaches have been proposed for email spam detection, including rule-based methods, content-based methods, and machine learning-based methods. Machine learning-based methods have gained popularity in recent years due to their ability to learn from large datasets and identify patterns and features that are characteristic of spam emails. Naive Bayes algorithm is one of the most used machine learning algorithms for email spam detection. It is a probabilistic algorithm that uses Bayes' theorem to calculate the probability of a given email being spam or not. In the Naive Bayes algorithm, each email is represented as a bag of words, and the probability of the email being spam is calculated based on the occurrence of words in the email. The algorithm assumes that the occurrence of words in an email is independent of each other, hence the name "Naive Bayes." One of the advantages of the Naive Bayes algorithm is its simplicity and efficiency. It can be trained on a large corpus of emails and can quickly classify new emails as spam or not. Additionally, the algorithm can handle a large number of features, making it suitable for high-dimensional data like text. In summary, email spam detection using Naive Bayes algorithm is a popular and effective method that has been widely used in research and industry. The algorithm's simplicity and efficiency make it a popular choice for email spam detection, and it has been shown to achieve high accuracy rates in identifying spam emails.

1.2 MOTIVATION

The motivation for email spam detection using Naive Bayes algorithm stems from the need to protect users from the risks associated with spam emails. Email is one of the most commonly used communication channels, and the volume of emails sent every day is enormous. A significant portion of these emails are spam emails, which are unwanted and often malicious. Spam emails can contain phishing links, malware, and fraudulent messages, which can lead to several problems, including identity theft, financial loss, and damage to reputation. Email service providers and users need an effective method to detect and filter out spam emails to improve the user experience and prevent security breaches. Machine learning algorithms, especially Naive Bayes, have been shown to be effective in email spam detection. Naive Bayes is a probabilistic algorithm that can be trained on a large corpus of emails to identify patterns and features that are characteristic of spam emails. The motivation for email spam detection using Naive Bayes algorithm is to provide an effective and efficient method for detecting and filtering out spam emails. The proposed approach can help in improving the overall user experience and reducing the risk of security breaches, which is crucial in today's digital world.

1.3 OBJECTIVE

Accurate classification: The system should accurately identify and classify emails as spam or ham with high precision and recall.

Minimizing false negatives: False negatives are legitimate emails that are incorrectly classified as spam. Minimizing false negatives is important because it ensures that important messages are not lost.

Minimizing false positives: False positives are spam emails that are incorrectly classified as legitimate. Minimizing false positives is important because it reduces the amount of spam that users receive.

Efficient processing: The system should be able to process many incoming emails quickly and efficiently.

Easy to use: The system should be easy for users to set up and use.

Robust to changes: The system should be robust to changes in spam tactics and be able to adapt to new forms of spam.

1.4 SCOPE OF THE PROJECT

The scope of the project for email spam detection using Naive Bayes algorithm involves designing and implementing a system that can effectively detect and filter out spam emails. The project will involve the following tasks:

1**. Data Collection**: Collecting a dataset of emails that contains both spam and non-spam emails.

2. **Data Preprocessing**: Preprocessing the email data to remove unnecessary information, such as email headers, and converting the emails into a suitable format for further analysis.

3. **Feature Extraction**: Extracting relevant features from the preprocessed email data that can be used to classify emails as spam or non-spam. The features could include word frequency, presence of specific words or phrases, and other characteristics.

4. **Training the Naive Bayes Model**: Training a Naive Bayes model on the extracted features to learn to classify emails as spam or non-spam.

5**. Model Evaluation**: Evaluating the performance of the trained model using metrics such as accuracy, precision, recall, and F1 score.

6. **Deployment**: Deploying the trained model in a system that can effectively filter out spam emails in real-time.

2. LITERATURE SURVEY

2.1 LITERATURE SURVEY

1. In the paper [1] "Email Spam Filtering: A Review", the authors provide a comprehensive review of different techniques used for email spam filtering, including Naive Bayes algorithm. The study highlights the effectiveness of Naive Bayes algorithm in email spam detection and discusses various features that can be used to improve its performance.

2.In the paper [2] "A Comparative Study of Machine Learning Techniques for Email Spam Filtering" compares the performance of different machine learning algorithms, including Naive Bayes, in email spam filtering. The authors conclude that Naive Bayes algorithm is one of the most effective algorithms for email spam filtering, with high accuracy and low false positive rates.

3. In the paper [3] "An Effective Email Spam Filtering System Using Machine Learning Techniques", the authors propose an email spam filtering system that uses Naive Bayes algorithm and other machine learning techniques. The study shows that Naive Bayes algorithm outperforms other machine learning algorithms in terms of accuracy and computational efficiency.

2.2 INFERENCE OF LITERTURE SURVEY

1. In the paper, proposed system attempts to use machine learning techniques to detect a pattern of repetitive keywords which are classified as spam. The system also proposes the classification of emails based on other various parameters contained in their structure such as Cc/Bcc, domain and header. Each parameter would be considered as a feature when applying it to the machine learning algorithm. The machine learning model will be a pre-trained model with a feedback mechanism to distinguish between a proper output and an ambiguous output. This method provides an alternative architecture by which a spam filter can be implemented. This paper also takes into consideration the email body with commonly used keywords and punctuations.

2.They describe a focused literature survey of Artificial Intelligence Revised (AI) and Machine learning methods for email spam detection. They have used the “image and textual dataset for the e-mail spam detection with the employment of various methods.

3.SYSTEM DESIGN

3.1 OVERALL ARCHITECTURE

The overall architecture of email spam detection using naive Bayes algorithm can be broken down into the following steps:

1.Data Preprocessing: The first step in building a spam detection model is to preprocess the data.

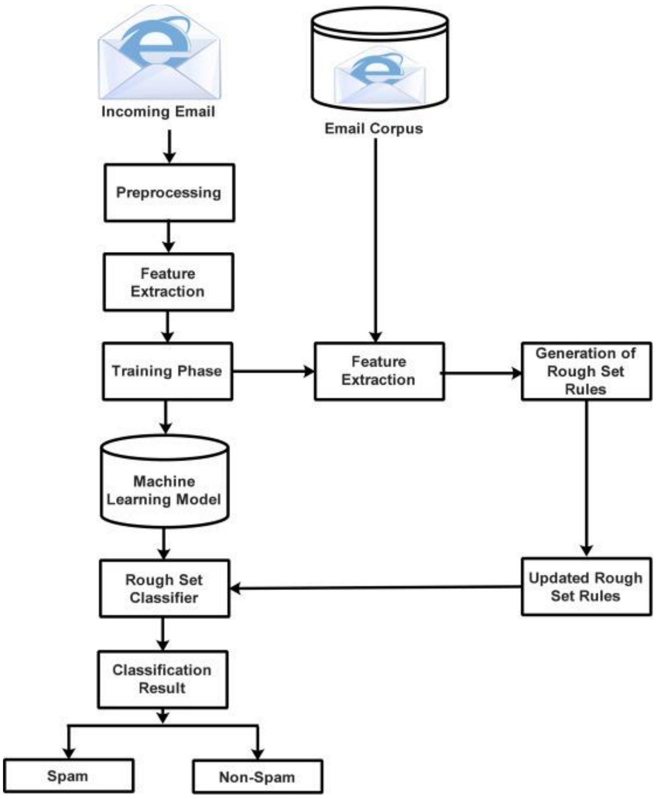
This involves cleaning and preparing the email data by removing unwanted characters, converting all letters to lowercase, and tokenizing the text.

2. Feature Extraction: The next step is to extract features from the preprocessed email data. The most common approach is to use the Bag-of-Words (Bow) model, which represents each email as a vector of word counts. Other feature extraction techniques include TF-IDF (Term Frequency-Inverse Document Frequency) and n-grams.

3. Training: Once the features have been extracted, the next step is to train the model. In the case of naive Bayes, the model is trained by estimating the probability of each feature occurring in spam and non-spam emails. This is done by calculating the conditional probability of each feature given the class (spam or non-spam).

4. Classification: After the model has been trained, it can be used to classify new emails as either spam or non-spam. This is done by calculating the probability of each email belonging to each class (spam or non-spam) using Bayes' theorem. The class with the highest probability is then assigned to the email.

5. Evaluation: The final step is to evaluate the performance of the model. This is typically done by splitting the data into training and testing sets and calculating metrics such as accuracy, precision, recall, and F1 score.



3.2 MODULE DESCRIPTION

3.2.1 DATA COLLECTION

The first step of Module description is collection of spam data. Get the spam data. Data is essential ingredient before we can develop any meaningful algorithm. . For this project We are collecting data set from the website called Kaggle. The data which is collected Consist of spam and ham data. Furthermore, in the ham data, they are easy and hard. Which mean, there is some non-spam data that has a very high similarity with spam data.

3.2.2 IDENTIFY THE SPAM EMAILS

There are some specific features to identity the received mail is spam / not. some of the features of identity the spam mail are.

1. If the received mail is asking personal information

2. If mail consist of urgent / Threatening messages in the subject

3. Spam mails mostly consist of typos/strange phrasing (sending mails from unauthorized website by making small change in the website name.)

4. If the received mail consists of large storage

5. Is the sender know to you and does the email address match the name?

3.2.3 IMPORTING LIBRARIES

The python libraries that are used in the code as follows:

NUMPY

NumPy is used for working with arrays. It also has functions for working in domain of linear algebra, Fourier transform, and matrices. NumPy stands for Numerical Python.

PANDAS

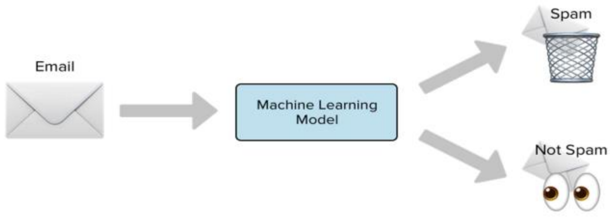
Pandas used for working with data sets. It has a function for analyzing, cleaning, exploring, and manipulating data. The name “Pandas” has a reference to both “Panel Data”, and “Python Data Analysis.”

SKLEARN

Scikit-learn (SK-learn) is the most useful and robust library for machine learning in python. It provides a selection of efficient tools for machine learning and statistical modelling including classification, regression, clustering, and dimensionality reduction via a consistence interface in Python.

3.2.4 CLASSIFICATION OF SPAM MAILS

In this all the received mails are classified into ham and spam mails. Ham consists of useful information whereas spam consist of unsolicited messages. The classification can be done by supervised learning technique. In this technique using a trained model. receiving mail can be labelled as spam or ham. If the mail is spam it is moved to spam folder



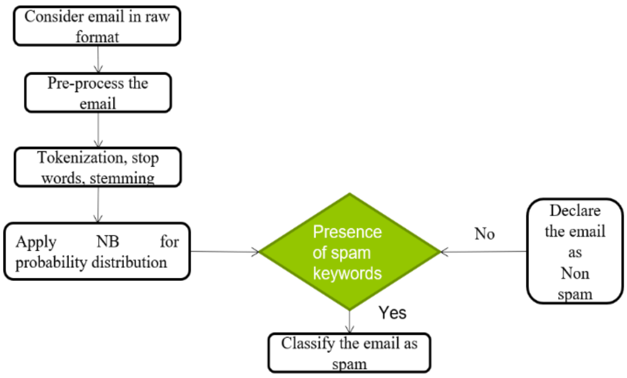
3.2.5 VERIFYING THE SPAM MAIL

Verifying the spam mail, we use naive bayes algorithm work by bayes classifier algorithm. Naive bayes classifier works by correlating the use of tokens (typically words, or sometimes other things) with spam and not spam emails and then using bayes theorem to calculate a probability that an email is or not spam . Sometimes informative mails are stored in spam folder. For classifying those we use naive bayes algorithm. It works by comparing the incoming mail with a dataset of emails, which are categorized into spam and ham.

3.2.6 DENY THE SPAM EMAILS

we are classifying into spam and ham mails. If the mail is found to be spam it enters spam folder. If the mail is found to be ham it deny from the spam folder. now if the mail present in spam folder is informative mail it moves to inbox deny from spam folder. if the mail present in spam folder is not informative mail allow spam folder.

3.3 FLOW CHART


Consider email in raw format
Pre-process the email
Tokenization, stop words, stemming
Apply NB for probability distribution
Presence of spam keywords
No
Declare the email as Non spam
Yes
Classify the email as spam

3.3.1 FLOW CHART EXPLANATION

STEP 1: First consider an email as a raw format.

STEP 2: Next Pre-Processing the given email

STEP3: Next we apply a spam detection steps or methods named as Tokenization, Stop Words, Stemming

STEP4: Now, Apply on Naive Bayes model for probability distribution.

STEP5: Now calculate the probability of each word appearing in spam and non-spam mail.

STEP6: Presence of spam keywords yes means its consider as “SPAM”, and presence of spam keyword no means “NON SPAM”

4. CONCLUSION

In conclusion, Naive Bayes is a powerful and effective algorithm for email spam detection. Its accuracy, efficiency, and scalability make it an ideal choice for addressing the challenges posed by spam emails in today's digital landscape. By leveraging probabilistic calculations and assuming independence between features, Naive Bayes can accurately classify incoming emails as spam or non-spam. Its simplicity and speed enable it to process large volumes of data in real-time, making it suitable for handling the ever-increasing influx of emails. Implementing Naive Bayes for email spam detection not only helps protect users from the nuisances of spam emails but also safeguards against potential security threats such as phishing attempts and malware distribution. By filtering out unwanted and potentially harmful content, Naive Bayes plays a crucial role in maintaining the integrity and security of communication channels. As the sophistication and prevalence of spam emails continue to grow, Naive Bayes remains a valuable tool in the fight against this issue. Its effectiveness, coupled with its ability to adapt and handle new types of spam, makes it a reliable solution for email providers, businesses, and individual users alike. In conclusion, Naive Bayes offers an efficient, accurate, and scalable approach to email spam detection, providing peace of mind and enhancing productivity in the digital realm.

ACKNOWLEDGMENT

We Would Like To Express Our Sincere gratitude to Mrs.M. Mahabooba, Professor For His Valuable Comments That Led To Substantial Improvement On Our Earlier Version Of This Manuscript

REFERENCES

[1] [2]. Nikhil kumar, sanket sonowal, nishant “email spam detection using machine learning algorithms” Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020) IEEE

[2] [3]. Simran Gibson, Biju Issac, Li Zhang, Seibu Mary Jacob “Detecting Spam email with machine Learning Optimized with bio-inspired metaheuristic algorithms” IEEE VOLUME 8, 2020

[3] [4]. Ganiev Salim Karimovich, Khamidov Sherzod Jaloddin ugli, Olimov Iskandar Salimbayevich “Analysis of machine leaning method for filtering spam messages in email services “ 2020 International Conference on Information Science and Communications Technologies (ICISCT) | IEEE nov 2020

[4] [5]. Mansoor RAZA and Nathali Dilshani Jayasinghe, Muhana Magboul Ali Muslam “A Comprehensive Review on email spam classification using Machine Learning Algorithms” 2021 International Conference on Information Networking (ICOIN) IEEE 02 February 2021

[5] [6]. Nandhini.S, DR.Jeen Marseline.K.S “Performance Evaluation of machine algorithms for email spam detection” 2020 International Trends in Information Technology and Engineering (ic-ETITE) IEEE 2020

[6] [7]. Mahammad Abdullahi, Abdulmalik D. Mohammed, Opeyemi O. Abisoye “A review on Machine Learning Techniques for images-based spam emails detection” Proceedings of the 2020 IEEE 2nd International Conference on Cyberspace (Cyber Nigeria) IEEE 2020

[7] [8]. Priya.S, Annie Uthra.R “An Effective Concept Drift Detection Technique with Kernel Extreme Learning Machine for Email Spam Filtering” Proceedings of the Third Sustainable Systems [ICISS 2020] IEEE 2020

[8] [9]. Fahima Hossain, Mohammed Nasir Uddin, Rajib Kumar Halder “Analysis of Optimized Machine Learning and Deep Learning Techniques for Spam Detection” International IOT, Electronics and Mechatronics Conference (IEMTRONICS) IEEE 2021

[9] [10]. Sunday Olusanya Olatunji “Extreme Learning Machines and Support Vector Machines Models for Email spam detection” Canadian Conference on Electrical and Computer Engineering (CCECE) IEEE 2017