INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 10, Issue 1 - V10I1-1287)
Available online at: https://www.ijariit.com

Preventia: Spam Alert System

Kaustubh Chaubey
kaustubhchaubey23@gmail.com
Shree L.R. Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Naufil Ahmed Siddique
naufil72@gmail.com
Shree L.R. Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Ayush Singh
ayush.27112@gmail.com
Shree L.R. Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Reena Kothari
reena.ostwal@slrtce.in
Shree L.R. Tiwari College of Engineering, Mira Bhayandar, Maharashtra

ABSTRACT

In the ever-evolving landscape of digital communication, the proliferation of spam content continues to be a pressing concern. The "Preventia: Spam Alert System" project aims to tackle this challenge head-on by implementing a comprehensive and intelligent solution for identifying and classifying spam. Through the application of the Naive Bayes algorithm, the system gains the ability to process vast amounts of data efficiently and make probabilistic predictions about whether incoming content is spam or legitimate. To bolster its capabilities further, the Spam Alert System integrates cutting edge technologies like natural language processing (NLP) to comprehend textual content more effectively. Subsequently, the Naive Bayes classifier evaluates the content, assigning a probability score that determines the likelihood of it being spam. The Spam Alert System's successful implementation delivers a powerful and flexible solution to the problem of spam, endowing users with heightened security and protection from fraudulent activities and potential privacy breaches. With seamless integration into popular communication platforms, users can enjoy real-time defence against malicious content across email, messaging services, and web browsers.

KEYWORDS: *Spam, Naive Bayes, NLP, Machine Learning, Fraud, Classifier.*

I. INTRODUCTION

In an era dominated by digital communication, the incessant surge of spam messages, emails, and calls has become a prevalent nuisance. As our reliance on various communication channels grows, so does the need for robust spam detection systems that can shield users from unwanted and potentially harmful content. This imperative has given rise to an advanced Spam Alert System, employing cutting-edge technologies to identify and thwart spam across multiple platforms.

The system incorporates Support Vector Machine (SVM) algorithms to tackle the ever-evolving landscape of SMS spam. SVM, a powerful machine learning technique, is adept at discerning patterns and classifying messages

based on their features. By leveraging SVM, our system analyses the characteristics of incoming SMS messages, swiftly distinguishing between legitimate and spam content with high accuracy.

Email, being a primary mode of communication, is frequently targeted by spam. To counter this, our Spam Alert System employs Naive Bayes classification for email spam identification. Naive Bayes excels in probabilistic reasoning, allowing the system to discern the likelihood of an email being spam based on various features. By adopting this approach, the system can efficiently sift through voluminous email traffic, providing users with a spam-free inbox experience.

Voice calls, too, have become a target for unsolicited communication. Our Spam Alert System employs Voice Signature Analysis coupled with behaviour-based Filtering to identify and block spam calls. Voice Signature Analysis scrutinizes unique voice patterns associated with known spam, while behaviour-based Filtering assesses call behaviour to identify anomalies indicative of spam activity. Together, these techniques fortify the system against unwanted and potentially fraudulent calls.

In conclusion, the Spam Alert System amalgamates state-of-the-art technologies to create a comprehensive defence against spam across SMS, email, and calls. By combining machine learning, probabilistic reasoning, and behavioural analysis, the system ensures a robust and adaptive shield against the ever-evolving landscape of digital spam, providing users with a secure and uninterrupted communication experience.

II. RELATED WORK

1) An Ensemble Learning Approach for SMS Spam Detection

- The paper discusses the challenges in SMS spam detection, such as the unbalanced proportion of spam and ham data and the extraction of efficient features from short messages.
- Various methods have been proposed to filter spam messages, but their accuracy still needs improvement.
- The paper proposes an ensemble learning method based on random forest and logistic regression algorithms to increase the accuracy of SMS spam detection[1] .
- The effectiveness of the proposed ensemble learning algorithm is evaluated based on accuracy and AUC (Area Under the Curve).
- The experimental evaluation shows that the proposed approach improves the accuracy of SMS spam detection, indicating its effectiveness in filtering spam messages.
- The proposed ensemble method outperforms random forest and logistic regression alone in terms of accuracy and AUC.
- Future work can focus on combining more algorithms, exploring different feature extraction techniques, and evaluating the generalization and robustness of the proposed ensemble learning method.

2) Enhancing the Naive Bayes Spam Filter

- Existing spam detection methods can be categorized into knowledge engineering and machine learning techniques[4] .
- Knowledge engineering methods rely on a set of rules to determine the legitimacy of emails, but they have disadvantages as they are based on spammers' previous methods and cannot adapt to changing spamming techniques.
- Machine learning methods, on the other hand, are customized based on the user and can adapt to changing spamming methods, but they are slower.
- Feature selection is an important issue in spam filtering as emails need to be converted into feature vectors, and incorrect classification can occur when spam and ham emails have the same feature vector.
- Most effective spam detection methods utilize some form of machine learning, such as Naive Bayes, Vector Space Models, clustering, neural networks, and rule-based classification.

3) Privacy Pro: Spam Calls Detection Using Voice Signature Analysis and Behaviour-based Filtering

- Voice spam detection technology differs greatly from junk email defence, and current voice spam avoidance techniques include list-based filtration, reputation-based filtration, Turing tests, and statistics-based filtration.
- List-based filtering detects spam calls by gathering user information and matching callers to black or white lists, but it is vulnerable to identity-based Sybil attacks and can be less effective if spammers modify their locations or identities[2] .

- Reputation-based filtration evaluates the whole user communication relationship and user evaluations of their peers to calculate reputation scores and determine whether a call is accepted or rejected. However, it relies on central reputation systems and may not be welcoming to new customers.
- Turing tests differentiate between human callers and automated voice spam but do not apply to human-initiated spam. They can be time-consuming and may not be suitable for all types of spam calls.
- Statistics-based filtering categorizes calls based on the similarity of the incoming caller's behaviour to the recipient's prior behaviour. However, it can be ineffective if the statistical characteristics of unwanted calls change.

4) Critical Analysis of Spam Detection Techniques

- E-mail spam continues to be a nuisance despite available anti-spam techniques.
- Signalling protocol techniques can enhance spam prevention by establishing genuine communication between sender and receiver servers.
- Content-based filters face challenges with obfuscated content and image spam.
- Bayesian filters have been effective in blocking spam while minimizing false positives and negatives.
- Whitelist and blacklist filters provide control over allowed and blocked email addresses.
- Challenge-response filters prompt users to validate their authenticity, reducing unwanted emails.
- Community filters allow users to report spam characteristics, leading to automatic blocking for other users.
- Grey listing rejects emails from unknown sources, but may delay legitimate mail.
- Rate Limiting limits the number of calls or messages per day to control spammers.
- Reputation filtering uses comprehensive information about the source to block spam and reduce false positives.

III. DATASET

The dataset is thoughtfully organized into two distinct categories: "ham" and "spam." The term "ham" designates legitimate messages, while "spam" refers to unsolicited and potentially malicious content. This clear classification facilitates the training and evaluation of machine learning models, enabling researchers to develop robust algorithms capable of accurately discerning between genuine and spam messages.

Table 1: Statistics of given data

CLASS	PERCENTAGE
HAM	87%
SPAM	13%

About Email Dataset

This is a csv file containing related information of 5172 randomly picked email files and their respective labels for spam or not-spam classification.
The csv file contains 5172 rows, each row for each email. There are 3002 columns. The first column indicates Email name. The name has been set with numbers and not recipients' name to protect privacy. The last column has the labels for prediction: 1 for spam, 0 for not spam. The remaining 3000 columns are the 3000 most common words in all the emails, after excluding the non-alphabetical characters/words. For each row, the count of each word(column) in that email(row) is stored in the respective cells. Thus, information regarding all 5172 emails is stored in a compact dataframe rather than as separate text files.

About SMS Dataset

The SMS Spam Collection is a set of SMS tagged messages that have been collected for SMS Spam research. It contains one set of SMS messages in English of 5,574 messages, tagged according being ham (legitimate) or spam.

IV. APPROACHES FOR SPAM ALERT SYSTEM MODEL

Naive Byes for Email

In this model, we are firstly creating a dictionary that includes a library named "stopwords" to remove all possible helping verbs from the content mentioned in the email. In the next phase, features are generating to train the dataset. After that our algorithm will be executed to check the possibility of an email to be spam or not. Finally, the machine learning model will be tested on a real-world emailing environment.

All paragraphs must be indented. All paragraphs must be justified, i.e. both left justified and right-justified.
Firstly, we have to prepare the data for generating a dictionary for our algorithm. This dictionary can be further utilized to extract the desired features which determine our algorithm. These are parameters through which a user can segregate spam emails and general or simply non-spam emails. To segregate junk or non-junk emails, first of all, we have collected some specific words and insert them into the dictionary which will be utilized in the proposed model. In the literature, numerous words extracting methods exists. So, there is a need to select the most suitable model precisely so we have used Naïve byes for Email Spam Detection.

Support Vector Machine for SMS

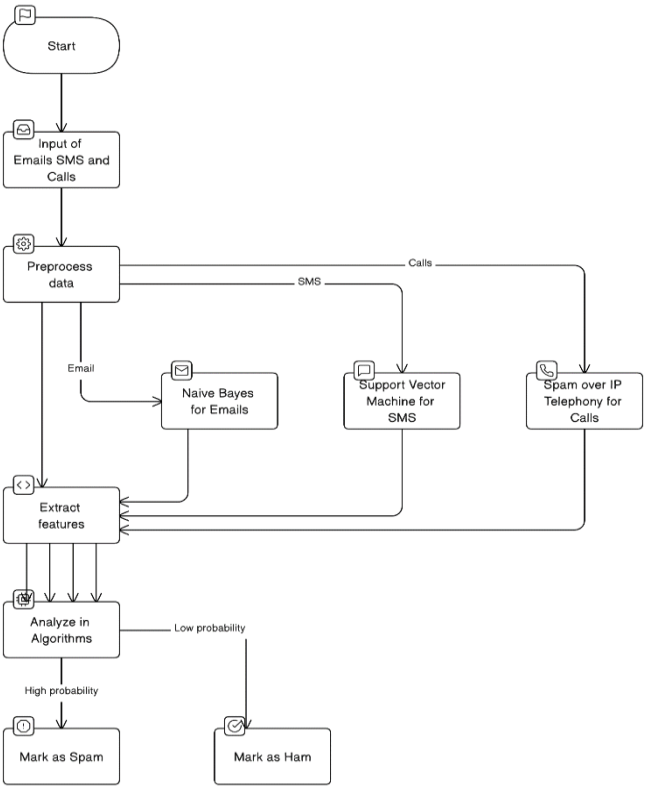
Support Vector Machines (SVM) stands as a cornerstone in machine learning, particularly renowned for its prowess in tackling classification quandaries. It's a robust tool adept at discerning patterns within datasets, especially in high-dimensional spaces.

At its core, SVM endeavours to unearth a hyperplane in the feature space that adeptly partitions data points into distinct categories. Visualizing this, in two-dimensional space, this hyperplane equates to a line, while in three-dimensional space, it transforms into a plane. The crux lies in maximizing the margin, the space between the hyperplane and the closest data points known as support vectors, hence the moniker "Support Vector Machines". In the realm of SMS spam detection, SVM proves invaluable. By identifying recurring keywords like 'cheap', 'sale', or 'needy', SVM learns to discern spam messages from legitimate ones. This process unfolds as a classification conundrum, where SVM excels due to its innate ability to process numeric data. Here, we leverage techniques like CountVectorizer to translate textual data into numeric form, ensuring SVM's efficacy. It's not just a matter of effectiveness, but also reliability. SVM often emerges as the optimal choice across diverse classification conundrums, underpinned by its capacity to handle nuanced datasets with finesse. In essence, SMS spam detection underscores a quintessential classification challenge, where SVM shines as a stalwart ally, navigating the labyrinth of textual data to discern patterns and safeguard against unsolicited messages.

Spam over telephony for calls

Spam over telephony, commonly known as "phone spam" or "robocalls," refers to unsolicited and often unwanted calls made to individuals or businesses for various purposes such as telemarketing, scams, or phishing attempts. Combatting phone spam poses a significant challenge due to its disruptive nature and potential to invade privacy.

V. FLOW DIAGRAM



VI. CONCLUSION

In conclusion, our Spam Alert System is a comprehensive defence mechanism designed to combat unsolicited communications across various platforms, including email, SMS, and telephony. Leveraging the power of machine learning algorithms, each component of our system contributes to the overarching goal of filtering out spam and preserving the integrity of digital communications. In summary, our Spam Alert System represents a holistic approach to combating unwanted communications, offering users a comprehensive defence against spam across multiple channels. Through the strategic integration of Naive Bayes, SVM, and advanced telephony spam detection technologies, we empower users to take control of their digital interactions and enjoy a safer and more enjoyable online experience.

VII. REFERENCES

[1] S. Hosseinpour and H. Shakibian, "An Ensemble Learning Approach for SMS Spam Detection," 2023 9th International Conference on Web Research (ICWR), Tehran, Iran, Islamic Republic of, 2023, pp. 125-128, doi: 10.1109/ICWR57742.2023.10139070.

[2] A. Kwong, J. H. Muzamal and Z. Khan, "Privacy Pro: Spam Calls Detection Using Voice Signature Analysis and Behavior-Based Filtering," 2022 17th International Conference on Emerging Technologies (ICET), Swabi, Pakistan, 2022, pp. 184-189, doi: 10.1109/ICET56601.2022.10004692.

[3] A. Alzahrani and D. B. Rawat, "Comparative Study of Machine Learning Algorithms for SMS Spam Detection," 2019 SoutheastCon, Huntsville, AL, USA, 2019, pp. 1-6, doi: 10.1109/SoutheastCon42311.2019.9020530.

[4] W. Peng, L. Huang, J. Jia and E. Ingram, "Enhancing the Naive Bayes Spam Filter Through Intelligent Text Modification Detection," 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), New York, NY, USA, 2018, pp. 849-854, doi: 10.1109/TrustCom/BigDataSE.2018.00122.

[5] R. J. Ben Chikha, T. Abbes and A. Bouhoula, "A SPIT detection algorithm based on user's call behavior," 2013 21st International Conference on Software, Telecommunications and Computer Networks - (SoftCOM 2013), Split, Croatia, 2013, pp. 1-5, doi: 10.1109/SoftCOM.2013.6671851.

[6] P. Wan and M. Uehara, "Spam Detection Using Sobel Operators and OCR," 2012 26th International Conference on Advanced Information Networking and Applications Workshops, Fukuoka, Japan, 2012, pp. 1017-1022, doi: 10.1109/WAINA.2012.24.

