iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.61187

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Enhancing Spam Filtering: A Machine Learning Approach to SMS and Email Fraud Classification

Mrs. T.Kavitha[1] , M Lakshmana Rao[2], Arava Harshavardhan[3], T S Sivamani Pallamraju[4], Rapaka Sasank[5]
[1]*Assistant Professor* , [2, 3, 4, 5] *B.tech Students Department of Information Technology, Pragati Engineering College, Surampalem , Andhra Pradesh, India*

Abstract: ***Spam is an unsolicited text message or SMS that may contain malicious content and is sent on mobile devices. Fraudsters send fictitious texts in an attempt to get victims to reply to their messages, and they may also steal account numbers, passwords, and other private information. It was suggested to use a model built around algorithms for machine learning to prevent falling for scammers' tricks. The Naïve Bayes method and term frequencies-inverse document frequency vectorizer are used to implement the suggested model. acquired the dataset from the Kaggle database and used it to train the model. The PyCharm IDE can be used to access the local host webpage that makes up this model. The obtained findings indicate a 95% accuracy and 100% precision for the model.***

Keywords: Spam SMS, Spam Email, Machine Learning, Naïve Bayes, Cyber Crime, Cyber Scam

I. INTRODUCTION

The entire world is becoming more digital. People communicate, transfer money, and engage in other activities that simplify life. It offers a lot of benefits, but it also has a lot of drawbacks. These days, people are readily duped by internet scammers that target them. Individuals may receive offers, unknown phone numbers, dubious links, and other things via social media, SMS, and email. Messages might be directed toward specific individuals or be received at random. Occasionally, communications may appear to be non-spam, deceiving people and maybe leading to scam success. Online scams fall under the category of cybercrime, for which the perpetrator may face legal consequences. However, because the public is unaware of these crimes, they may go unreported, which could encourage additional scam operations. The public is warned by banks, telecom companies, and cybercrime agencies about attackers and spammers that trick people via texts, links, and emails. However, the main reason why people fall victim to cyberscams is that they usually don't know if the emails and messages they get are real or fake [19]. In the last three years, 42% of Indians claimed being victim to financial fraud, and 74% of those surveyed said they were unable to get their money back. This information was based on a poll conducted by a private organization named Local Circles. To tackle these cyberscams, an algorithm based on machine learning has been proposed to help individuals determine whether the emails and texts they are receiving are spam. The user is able to paste the statement into the open-source website established whenever they believe it to be harmful.

II. LITERATURE SURVEY

Compared to other forms of detection, Lutfun Nahar Lota et al. [1] described how SMS identification of spam can be more difficult. The variety of algorithms available and the room for creativity. I was able to examine a lot of additional research on the exact same topic thanks to this work. Paras Sethi et al. [2] expounded on the extent of SMS spam's global concern. many algorithms that are available for analysing the model, determining which of them is the best, and researching various filtering techniques. In their summary, Shafi'i Muhammad Abdulhamid1 et al. [3] addressed the approaches that are commonly employed, the difficulties that may arise, and opportunities for further study regarding filtering out spam and spam detection in mobile SMS. In order to verify various model parameters, such as correctness, this study detailed the possible filtering strategies. M. Rubin Julis et al. [4] described the many stages of data extraction, how to represent and evaluate the dataset, and a number of filtering methods. In order to show the correctness and precision of every model used for analysis, Amani Alzahrani et al. [5] described how to present the cleaned information by changing the titles of the columns and how to represent the analyzed data using various graphs for various parameters. Along with the analysis, Naina Nisar et al. [6] went into detail on a variety of classifiers, including SVM, NB, k-NN, DT, RF, and Adaboost. examined the Voting Ensemble, which is used to test the model's correctness and precision throughout the training process. Sakshi Agarwal et al. [7] suggested including Indian spam messages in the dataset that was already gathered through worldwide data collection, analysis, and search. This work has examined several SMS spam detection techniques and stages, as well as numerous classifiers that can be applied to model training.

Opinion review spam was described by Michael Crawford et al. [8]. Here, spammers post fictitious reviews in an attempt to influence people. They post phony, fraudulent, and misleading reviews in an effort to win people over. suggested approach that makes use of natural language processing (NLP). believes that people should be aware that not all reviews are authentic and trustworthy, and that it's necessary to build a method for identifying and evaluating spam. Anju Radhakrishnan et al.'s summary of various machine learning methods for identifying spam emails using various email or SMS datasets was published in [9]. Nikhil Govil et al. [10] provided a detailed explanation of the processes taken to develop the model. The creation and evaluation of the Naïve Bayes algorithms, spam filtering algorithm, spam recognition techniques, and machine learning model are all covered in this work. Since SMS spam detection is a relatively new field and the techniques are still in the development stage, Shirani-Mehr et al. [11] provided an explanation there might not be enough trustworthy reviews. explained in detail how they looked at several literature studies and ran through various algorithms to arrive at a highly accurate solution. According to Fette, I. et al. [12], the internet has grown to be an essential benchmark in terms of information and knowledge security. It is employed to detect malevolent users and hackers. described the use of emails as phishing techniques. In order to improve performance in categorizing spam and ham SMS, Shaufiah et al. [13] combined the FPgrowth algorithm with the Naïve Bayes classifier. They explained that achieving accuracy is the most difficult aspect of the SMS spam filtering process. et al., Andronicus A. [14] outlined the primary use of random forest algorithms in an application of email fraud detection and improved prediction accuracy through training on a dataset of 2000 phishing and ham emails. Their final classification accuracy was 99.7%. The suggested paradigm for computational analysis uses machine learning methods and natural language processing, as explained by G. Tripathi et al. [15]. Here, the SVM and Naïve Bayes classifiers' behaviours are seen.

III. SYSTEM ANALYSIS

A. *Existing System*

The existing system for your project "Spam SMS (or) Email Detection and Classification using Machine Learning" is designed to tackle the issue of identifying and categorizing spam SMS messages to protect users from potentially harmful content and fraudulent activities. It employs the Naïve Bayes algorithm for classification and a term frequency-inverse document frequency (TF-IDF) vectorizer for feature engineering. The project utilizes a dataset obtained from Kaggle for training the model. The system includes a user-friendly interface in the form of a local host website, which allows users to input SMS messages and receive the classification results. Based on the abstract, the system has demonstrated strong performance with a 95% accuracy rate and a precision of 100%, indicating a high level of accuracy in correctly identifying spam messages and minimizing false positives.

DISADVANTAGES OF THE EXISTING SYSTEM

1) *Limited to Text-Based Messages:* The system primarily focuses on detecting and classifying text-based SMS messages. It may not be effective in identifying spam in other formats, such as multimedia messages or emails.
2) *Dependency on Training Data:* The system's performance heavily depends on the quality and representativeness of the training dataset from Kaggle. If the dataset is not diverse or up to date, it may not perform well on real-world spam messages.
3) *Overfitting Concerns:* Achieving 100% precision in the model could indicate overfitting to the training data, which may result in reduced performance on unseen data. It's essential to balance precision with other metrics and ensure the model generalizes well.
4) *Lack of Real-Time Updates:* The system may not have the capability to update its spam detection rules and algorithms in real time. Spam patterns can change over time, and the system may become less effective if it cannot adapt to new spamming techniques.
5) *Scalability and Deployment:* While the system is implemented locally through PyCharm IDE, deploying it at scale in a production environment may pose challenges. Ensuring the system can handle a large volume of SMS messages and maintaining its performance can be complex.

B. *Proposed System*

A proposed system for enhancing the existing "Spam SMS (or) Email Detection and Classification using Machine Learning" project could encompass several improvements and features.

1) *Multi-Modal Content Detection:* The proposed system would expand its capabilities beyond text-based messages, incorporating the ability to detect spam in multimedia messages, such as images, audio, and videos, as well as email content. This enhancement ensures comprehensive protection against a wider range of spam content.
2) *Dynamic Data Sources:* Instead of relying solely on a static dataset from Kaggle, the system could incorporate dynamic data sources to continuously update its spam detection algorithms. This might involve real-time data feeds, user-generated reports, or integration with external threat intelligence services to stay up to date with emerging spam patterns.
3) *Advanced Machine Learning Techniques:* In addition to Naïve Bayes, the proposed system could explore more advanced machine learning and natural language processing techniques, such as deep learning models (e.g., neural networks), ensemble methods, and topic modeling. This would potentially improve accuracy and adaptability to evolving spam tactics.
4) *Real-Time Updates:* The system should be designed to receive real-time updates and model retraining to stay ahead of evolving spam tactics. Continuous learning and adaptation ensure that it remains effective in identifying new spam threats as they emerge.

IV. SYSTEM DESIGN

SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture.

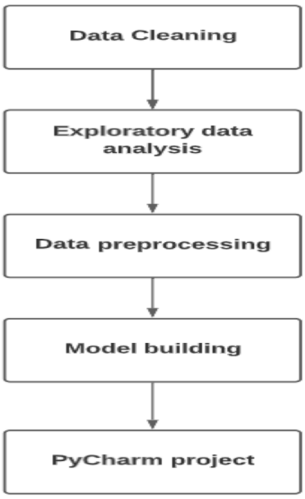


Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

MODULES

1) *Data Collection and Preprocessing Module:* This module is responsible for collecting data from various sources, such as SMS messages, multimedia messages, and emails. It preprocesses the data, including text normalization, removal of noise, and feature extraction. It ensures that the data is ready for analysis and classification.
2) *Machine Learning Model Module:* This module involves the implementation of machine learning models for spam detection and classification. It encompasses model training, validation, and evaluation. The module can include a variety of models, including Naïve Bayes, deep learning, and ensemble methods.
3) *Real-Time Threat Intelligence Module:* To stay updated with emerging spam tactics, this module continuously monitors and collects data from real-time threat intelligence sources, external APIs, and user-generated reports. It incorporates this information into the system to enhance its accuracy and effectiveness.
4) *User Interface and Reporting Module:* This module provides a user-friendly interface for users to interact with the system. Users can input messages, view classification results, and report false positives or negatives. It also generates reports and visualizations to convey the system's performance to users.
5) *Scalability and Deployment Module:* To ensure that the system can handle increased user demand and message volume, this module focuses on scalability and cloud deployment. It manages system resources, load balancing, and scalability mechanisms to ensure a seamless user experience.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

VI. RESULTS AND DISCUSSION

The spam SMS filtering approach that is offered is analysed using multiple methods, shown through charts and graphs, and evaluated based on efficiency, accuracy, and precision. It uses the Naïve Bayes classifier to implement TF-IDF. The suggested model is a webpage with a message writing block and an analysis button that indicates whether or not the message is spam.

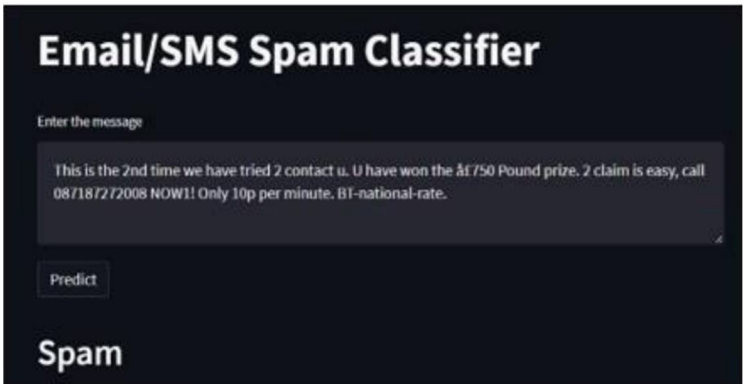


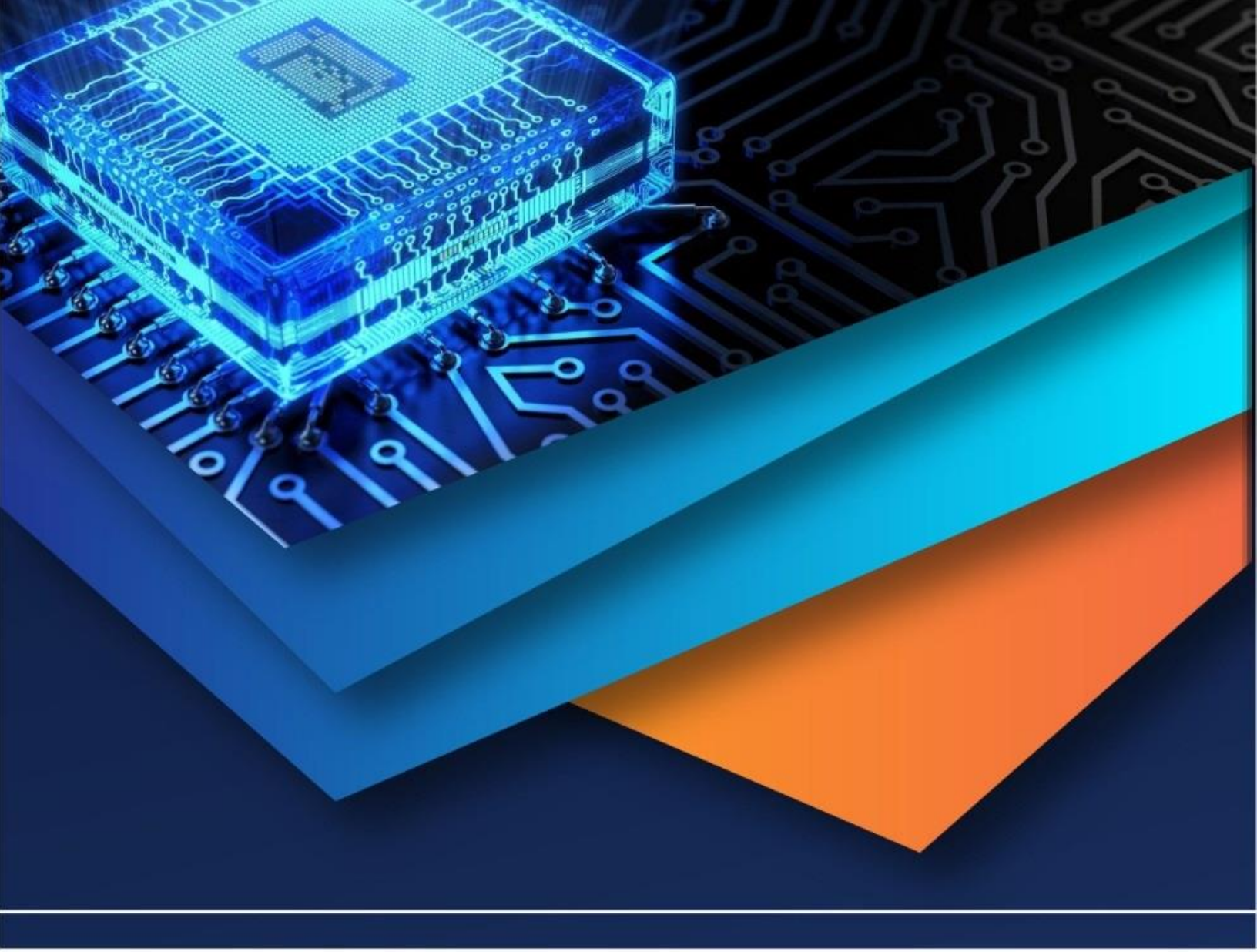
Fig 2. Based on the Context analysing SMS or email is Spam or not

VII. CONCLUSION AD FUTURE WORK

Because more people have access to internet access and mobile connectivity, the threat posed by spam SMS is growing rapidly on a global scale. India is experiencing a greater level of exposure to this phenomenon due to the cheaper cost of SMS services. To prevent fraudulent activities, the model suggests a machine learning-based remedy as a precaution. The suggested spam SMS filtering technique applies TF-IDF using the Naïve Bayes classifications and is evaluated using a variety of algorithms, graphs, and charts, and efficiency, accuracy, and precision at the end. The suggested model is a webpage with a message writing block and an analysis button that indicates whether or not the message is spam. Because of this, the model is user-friendly and suitable for individuals of all ages. We can safeguard ourselves against the majority of online scams because this model provides precision as well as accuracy of above 95%.

REFERENCES

[1] Lutfun Nahar Lota et al. “A Systematic Literature Review on SMS Spam Detection Techniques”, I.J. Information Technology and Computer Science, 2017, 7, 42-50.
[2] P. Sethi et al. “SMS spam detection and comparison of various machine learning algorithms,” International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), 2017, pp. 28-31.
[3] S. M. Abdulhamid et al. “A Review on Mobile SMS Spam Filtering Techniques,” IEEE Access, vol. 5, pp. 15650-15666, 2017.
[4] M.Rubin Julis et al. “Spam Detection in SMS Using Machine Learning Through Text Mining”, International journal of scientific & technology research, vol 9, Issue 02, 2020.
[5] A. Alzahrani et al. “Comparative Study of Machine Learning Algorithms for SMS Spam Detection,” SoutheastCon, 2019, pp. 1-6.
[6] N. Nisar et al. “Voting-Ensemble Classification for Email Spam Detection,” International Conference on Communication information and Computing Technology (ICCICT), 2021, pp. 1-6.
[7] S. Agarwal et al. “SMS spam detection for Indian messages,” International Conference on Next Generation Computing Technologies (NGCT), 2015, pp. 634-638.
[8] Michael Crawford et al. “Survey of Review spam detection using machine learning techniques”, Journal of Big Data, 2015.
[9] Anju Radhakrishnan et al. “Email Classification using Machine learning algorithms”, International Journal of Engineering and Technology (IJET), 2017.pp.335-340.
[10] N. Govil et al. “A Machine Learning based Spam Detection Mechanism,” International Conference on Computing Methodologies and Communication (ICCMC), 2020, pp. 954-957.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY