IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning-Based Fraud SMS or Email Identification and Categorization

1. lasya pithani 2 K. Chinna Nagaraju, 3 V. Anil Santhosh

1. M.Tech Scholar and student of C.S.E(AI), International School of Technology and Sciences for Women(Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh, lasya216w1daf13@gmail.com

2. Associate Professor of C.S.E., International School of Technology and Sciences for Women(Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh.

3. Professor and HOD of C.S.E., International School of Technology and Sciences for Women(Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh.

Abstract:

Spam is an unwanted message or SMS sent on mobile phones whose content may be malicious. Scammers send fake text messages to trick people into responding to their SMS and they may hack personal information, password, account number, etc. To avoid being tricked by scammers, proposed a model based on Machine learning Algorithms. The proposed model is implemented using the Naïve Bayes algorithm and term frequency-inverse document frequency vectorizer. Obtained the dataset from Kaggle and trained the model using it. This model consists of a local host website which is obtained through PyCharm IDE. Obtained results show that the model accuracy of 95% and a precision of 100%.

Key words: Machine Learning Techniques, Fraud SMS or Email Identification, Categorization

I. INTRODUCTION

Whole world is moving towards digitalization. People converse, send money and do many activities which make life easier. Even though it has many pros it also has many cons too. Nowadays people are targeted by online scammers and get tricked easily. People may receive suspicious links, unrecognized contact numbers, offers, etc. through emails, SMS, and social media. The messages can be received randomly or targeted on particulars. Sometimes the messages might seem to be non-spam which can trick people and can get successful in scamming. Online scams come under cybercrime and the thief can be sentenced to punishments but due to a lack of awareness in the general public, these crimes can go unnoticed which may promote these scam activities more. Cybercrime offices, telecom companies, and banks warn people about spammers and hackers who trick people by sending messages, links, and emails. But normally people are not aware of whether the messages and emails they get are verified or fake due to this reason cyber scams happen [19]. A private firm named Local Circles conducted a survey in which the statistics showed that in the last 3 years 42% of Indians faced financial fraud and 74% of people failed to retrieve the money. To overcome these cyber scams, proposed a model based on machine learning which helps individuals to check if the messages and emails they are receiving are spam or not. Whenever the user feels the message is unsafe, they can copy and paste it into the opensource site created.

II. LITERATURE SURVEY

Spam SMSs are unsolicited messages to users, which are disturbing and sometimes harmful. There are a lot of survey papers available on email spam detection techniques. But SMS spam detection is comparatively a new area and systematic

literature review on this area is insufficient. In this paper, we perform a systematic literature review on SMS spam detection techniques. For that purpose, we consider the available published research works from 2006 to 2016. We choose 17 papers for our study and reviewed their used techniques, approaches and algorithms, their advantages and disadvantages, evaluation measures, discussion on datasets and finally result comparison of the studies. Although, the SMS spam detection techniques are more challenging than email spam detection techniques because of the regional contents, use of abbreviated words, unfortunately none of the existing research addresses these challenges. There is a huge scope of future research in this area and this survey can act as a reference point for the future direction of research.

Past few years have seen increase in the number of spam emails and messages. Legal, economic and technical measures can be used to tackle spam SMS's nowadays. A key role is being played by Bayesian filters in stopping this problem. In this paper, we analyzed and studied the relative strengths of various machine learning algorithms in order to detect spam messages which are sent on mobile devices. We have acquired the data from on open public dataset and prepared two datasets for our testing and validation purposes. Accuracy in detecting spam messages was the first priority in ranking these algorithms. Our results clearly demonstrate that different machine learning algorithms under different features tend to perform differently in classifying spam messages.

Under short messaging service (SMS) spam is understood the unsolicited or undesired messages received on mobile phones. These SMS spams constitute a veritable nuisance to the mobile subscribers. This marketing practice also worries service providers in view of the fact that it upsets their clients or even causes them lose subscribers. By way of mitigating this practice, researchers have proposed several solutions for the detection and filtering of SMS spams. In this paper, we present a review of the currently available methods, challenges, and future research

directions on spam detection techniques, filtering, and mitigation of mobile SMS spams. The existing research literature is critically reviewed and analyzed. The most popular techniques for SMS spam detection, filtering, and mitigation are compared, including the used data sets, their findings, and limitations, and the future research directions are discussed. This review is designed to assist expert researchers to identify open areas that need further improvement.

The development of the cell phone clients has prompted a sensational increment in SMS spam messages. Despite the fact that in many parts of the world, versatile informing channel is right now viewed as "spotless" and trusted, on the complexity ongoing reports obviously show that the volume of cell phone spam is drastically expanding step by step. It is a developing mishap particularly in the Middle East and Asia. SMS spam separating is a similarly late errand to arrangement such an issue. It acquires numerous worries and convenient solutions from SMS spam separating. Anyway it fronts its own specific issues and issues. This paper moves to deal with the undertaking of sifting versatile messages as Ham or Spam for the Indian Users by adding Indian messages to the overall accessible SMS dataset. The paper examinations distinctive machine learning classifiers on vast corpus of SMS messages for individuals.

The short message service (SMS) became popular after it was initially provided as a service in the second-generation (2G) terrestrial mobile network architecture (Global System for Mobile Communication - GSM). Its popularity has been exploited by some advertising companies and others to spread unwanted advertising, communicate advertising offers, and send unwanted material to the end users. These undesirable messages, known as spam, make it difficult for the users to receive the desirable messages and make them frustration and irritation. Consequently, there are measures that various experts have implemented in filtering out these spam messages and blocking them from reaching the end users. Most of the solutions have followed the success of email spam filtering and utilized machine learning techniques to filter spam messages. The popular machine learning techniques that have successfully been used include logistical regression, Naïve Bayes algorithms, Support Vector Machine (SVM), and neural networks. The present study adopts these techniques in filtering spam messages and measures their accuracy to determine the most effective method of filtering spam messages. Based on the findings, the neural network performs best as the trained classifier model used to classify incoming messages as ham or spam.

III. SYSTEM ANALYSIS

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.

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 email

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.

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

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.

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.

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 modelling. This would potentially improve accuracy and adaptability to evolving spam tactics.

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.

Multi-Modal Protection: The system's ability to detect and classify various types of spam, including text, multimedia, and email-based spam, provides comprehensive protection to users across different communication channels, reducing the risk of exposure to diverse spam threats.

Improved Accuracy: By incorporating advanced machine learning techniques and continuously updating its algorithms, the system can achieve higher accuracy rates in identifying and filtering out spam messages. This leads to a reduced likelihood of false positives and an enhanced user experience.

Real-Time Threat Response: The integration of real-time updates and dynamic data sources enables the system to respond promptly to emerging spam patterns and threats. This real-time threat response is crucial for staying ahead of scammers and adapting to changing tactics.

Enhanced Scalability: Architecting the system for scalability and cloud deployment ensures that it can efficiently handle a growing user base and an increasing volume of messages. This scalability is essential to accommodate user needs and maintain system performance.

User-Friendly Experience: A well-designed user interface and a system that continually evolves to protect users from spam contribute to an overall user-friendly experience. Users can trust the system to keep their communications safe, leading to increased satisfaction and trust in the platform.

IV.SYSTEM DESIGN

SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture.

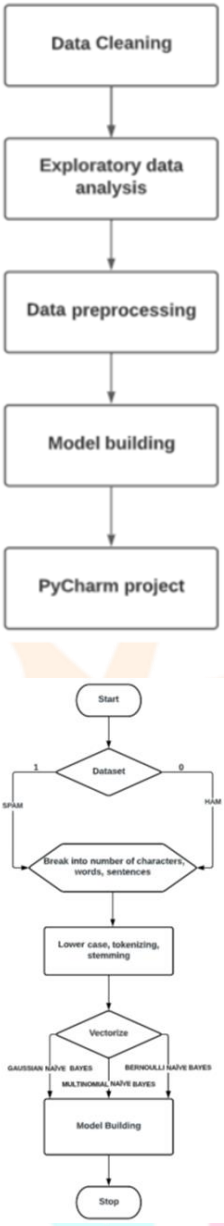


Fig 1. Methodology followed for proposed model

V. SYSTEM IMPLEMENTATION

MODULES

Data Collection and Pre-processing Module:

This module is responsible for collecting data from various sources, such as SMS messages, multimedia messages, and emails. It pre-processes the data, including text normalization, removal of noise, and feature extraction. It ensures that the data is ready for analysis and classification.

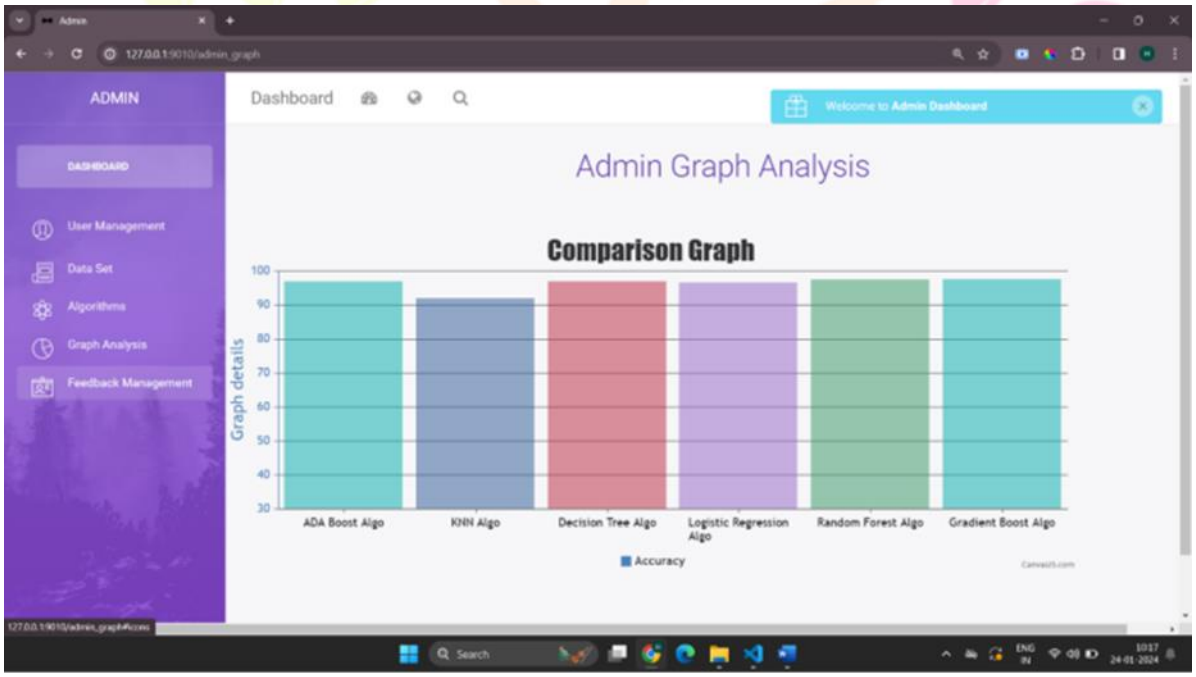
Machine Learning Model Module:
This module involves the implementation of machine learning models for spam detection and classification. It encompasses model training, validation, and evaluation. The module can include a variety of models, including Naïve Bayes, deep learning, and ensemble methods.

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.

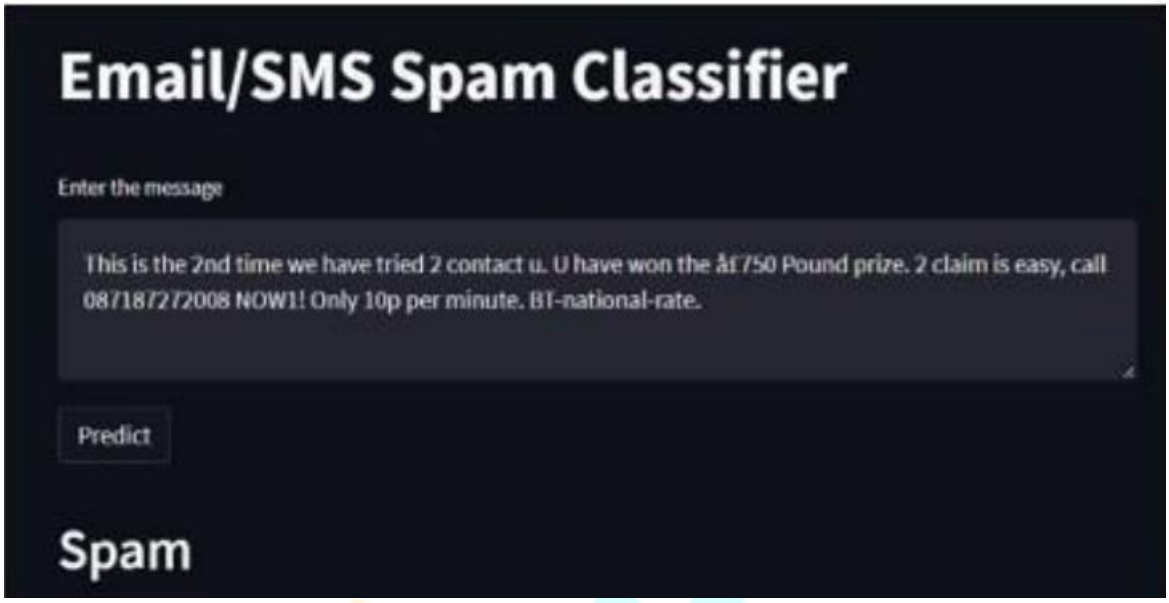
User Interface and Reporting Module:
This module provides a user-friendly interface for users to interact with the system. Users can input messages, view classification results, and report false positives or negatives. It also generates reports and visualizations to convey the system's performance to users.

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.• Regression metrics such as mean squared errors (MSE) or mean absolute errors (MAE) can be used to predict earnings.

VI. RESULTS AND DISCUSSIONS



In above diagram a discribes about verious algorithmic accuracy comparion graph



The presented spam SMS filtering method is analysed based on various algorithms, visualized through graphs and charts, and finally based on performance, accuracy, and precision; it implements TF-IDF with the Naïve Bayes classification. The proposed model is a website consisting of a block to write a message and a prediction button that informs whether the message is spam or not.

VI I. CONCLUSION AND FUTUREWORK

The danger of spam SMS is increasing all over the world at a very high rate and keeps on accelerating since access to the internet and mobile connectivity has increased. India is getting higher exposure to this phenomenon because of the availability of SMS services at lower cost. As a matter of precaution and to avoid fraud occurrences the model proposes a machine learning-based solution. The presented spam SMS filtering method is analysed based on various algorithms, visualized through graphs and charts, and finally based on performance, accuracy, and precision; it implements TF-IDF with the Naïve Bayes classification. The proposed model is a website consisting of a block to write a message and a prediction button that informs whether the message is spam or not. This makes the model easy to use and adaptable for all age groups of people. As this model gives accuracy and precision of more than 95%, to protect ourselves from most online scams.

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.

Biography of authors:

Lasya pithani was a MTech Scholar and student of student of C.S.E(Artificial Intelligence)., International School of Technology and Sciences for Women(Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh. **lasya pithani** is a dedicated research scholar specializing in Data Since ,Python and Machine Learning (ML), focusing on innovative approaches to solve complex real-world problems. With a strong academic foundation and a passion for computational technologies.

Mr K Chinna Nagaraju (Ph.D.) was an Associate Professor of C.S.E., International School of Technology and Sciences for Women (Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh. **Chinna Nagaraju** is a dedicated research scholar specializing in Artificial Intelligence (AI) and Machine Learning (ML), focusing on innovative approaches to solve complex real-world problems. Their research interests include developing advanced algorithms for predictive modelling, integrating hybrid ML-DL frameworks, and exploring the ethical and societal impacts of AI systems.

V Anil Santhosh was an Assistant Professor and HOD of C.S.E., International School of Technology and Sciences for Women(Autonomous), East Gonagudem, Rajanagaram–Andhra Pradesh. V Anil Santhosh is a dedicated research scholar specializing in Artificial Intelligence (AI) and Machine Learning (ML), focusing on innovative approaches to solve complex real-world problems. Their research interests include developing advanced algorithms for predictive modelling, integrating hybrid ML-DL frameworks, and exploring the ethical and societal impacts of AI systems. Their work primarily focuses on applications in renewable energy forecasting, natural language processing, and computer vision.