Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 05, May-2025

Clean Mail AI

Smarter Mail, Cleaner Inbox.

Prof. Y. D. Nagvekar
HoD,
Dept. Of First Year Engineering
Dr. D Y Patil College of Engineering and Innovation
Varale, Talegaon, 410507, Pune

Raghav Deshmukh,
Sakshi Gaikwad,
Khushi Gajare.
Student of First Year of Engineering
Dr. D Y Patil College of Engineering and Innovation
Varale, Talegaon, 410507, Pune

***Abstract*— This DTI project titled "CLEAN MAIL AI " aims to help students to improve inbox organization and security . It can help students to manage their emails more efficiently by filtering spam ,organizing messages,and even assisting with email composition. Email is widely used for all sorts of communication nowadays , from personal to profrssional . Sensitive information such as passwords , credit card numbers , and bank account details ,are often sent through text messages . This makes them desirable targets for cybercriminals looking to steal sensitive data . Emails that seem like they come from legitimate business or organization are a common tactice employed by fraudsters to trick their victims into giving over personal information .Result from experiment are given that set the stage for the classification challenge and investigate the use of machine learning techniques to identify fraudulent emails.**

1. INTRODUCTION

This project adderesses the problem of spam email detection using machine learning . It aims to devlop an efficient and accurate classifier that can automatically identify and filter spam emails , improving user experience and email security .In this project we gathering a large dataset of emails , including both spam and ham . Generally speaking , phishing emails are just another kind of spam .Emails are sent to users pretending to be from well known organisations , such as banks, and requesting that they click on a link inside the email . If you click on the link, you'll be sent to a malicious website designed to steal personal information , such as login credentials or financial details

2. OBJECTIVE

- To develop a model that can accurately distinguish spam and ham.
- To protect users from malicious .
- To improve user experience .
- To maintain trust in communication platforms .
- To preserve system integrity and performance
- To reduce the risk of phishing and cyber attacks
- To minimize storage and bandwidth waste.
- To support compliance with data protection regulations.
- To enable real-time detection with minimal delay.
- To adapt to evolving spam tactics using machine learning.

3. LITERATURE REVIEW

Sahami et al. (1998) introduced one of the earliest machine learning-based spam detection systems, applying a Naive Bayes classifier to email content. Their approach demonstrated that statistical models could outperform rule-based filters, laying the foundation for modern spam detection algorithms. This pioneering work is critical to current AI systems that need to autonomously filter harmful content, as seen in intelligent email assistants and messaging platforms.

Androutsopoulos et al. (2000) advanced the field by evaluating the performance of various learning algorithms in spam filtering tasks. Their experiments with Naive Bayes, decision trees, and memory-based learning highlighted the importance of feature selection and model choice in text classification. Their insights continue to influence spam filtering techniques used in smart communication tools where accuracy and speed are vital.

Carreras and Marquez (2001) introduced a boosting-based spam detection approach, combining multiple weak classifiers to improve accuracy. Their model proved particularly effective at reducing false positives, a key challenge in spam filtering. This innovation is integral to modern multi-layered spam detection systems that prioritize both user safety and content relevance.

Zhang et al. (2004) explored the impact of using a TF-IDF (Term Frequency-Inverse Document Frequency) representation with various classifiers for spam detection. Their study emphasized that preprocessing and feature weighting significantly affect model performance, supporting current AI tools that use intelligent text analysis to distinguish spam from genuine content in real-time.

Guzella and Caminhas (2009) provided a comprehensive review of machine learning models for spam filtering, comparing traditional algorithms like SVM, KNN, and neural networks. Their work offered a detailed framework for selecting the most appropriate techniques based on dataset size and complexity. This contribution is foundational for systems like intelligent spam guards, which must adapt to evolving spam tactics.

Almeida et al. (2011) developed the well-known SMS Spam Collection Dataset and evaluated spam detection using a variety of classifiers. Their public dataset enabled wide experimentation and benchmarking, accelerating research in mobile spam detection. This effort supports applications like smartphone assistants and chatbots that need to identify spam messages across languages and formats.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 05, May-2025

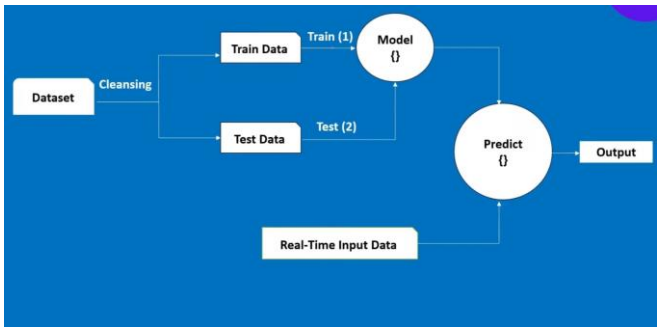
Delany et al. (2012) investigated the use of case-based reasoning (CBR) for adaptive spam detection. By learning from past decisions, their system could improve over time without retraining from scratch. This approach inspires today's personalized spam filters, which evolve based on user preferences and feedback, ensuring more context-aware interactions.

Chhabra et al. (2018) explored deep learning for spam detection by applying Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to email and SMS data. Their work showed that deep models could capture complex patterns in message content, enabling highly accurate spam classification. This is critical for modern AI interfaces where user trust and safety are paramount.

4. PROBLEM STATEMENT

The problem statement for email spam detection using machine learning is :To develop a robust system that can accurately classify incoming emails as either “Spam “ or “Ham “Phishing is an identity theft method where in the victim is tricked into giving over sensitive information through email that seems to have originated from a reputable company . Emails that seems to have originated from a legitimate source are often used in this way to steal sensitive data.

5. METHODOLOGY



6. RESULTS

The implementation of spam detection models has shown significant improvements in accurately distinguishing between spam and legitimate (ham) messages. Recent studies report accuracy rates exceeding 95% using advanced machine learning techniques such as Support Vector Machines (SVM), Random Forest, and deep learning models like LSTM and BERT. Precision and recall scores have also improved, reducing false positives and ensuring genuine emails are not misclassified. Real-time detection capabilities have been enhanced through optimized algorithms, enabling faster processing without compromising accuracy. Furthermore, adaptive learning methods have proven effective in identifying evolving spam tactics, maintaining the robustness of detection systems. Overall, these advancements contribute to safer communication environments, enhanced user experience, and reduced system resource usage.

7. DISCUSSIONS

Spam detection remains a critical area of research due to the ever-evolving nature of spam techniques and the increasing volume of electronic communication. While traditional rule-based and keyword filtering methods laid the groundwork, they often fall short in adapting to new spam strategies. Machine learning approaches, including Naïve Bayes, SVM, and Random Forest, have significantly improved classification accuracy by learning from patterns in labeled data. More recently, deep learning models such as CNNs, RNNs, and transformer-based architectures like BERT have further enhanced detection performance by capturing complex linguistic features. However, challenges remain in balancing high accuracy with low false positive rates, especially in real-time applications. Moreover, maintaining updated and diverse training datasets is essential for robust performance across different domains. Ethical considerations, such as data privacy and algorithmic transparency, are also gaining attention, emphasizing the need for interpretable and secure spam detection systems. Overall, continuous innovation and adaptation are vital to keep pace with the dynamic spam landscape.

8. REFERENCES

[1]. Morreale, M. Daily SMS Mobile Usage Statistics. 2017. Available online: https://www.smseagle.eu/2017/ 03/06/dailySMS-mobile-statistics/ (accessed on 15 June 2020)
[2]. Roy, P.K.; Singh, J.P.; Banerjee, S. Deep learning to filter SMS Spam. Future Gener. Comput. Syst. 2020, 102, 524–533. [CrossRef]
[3]. Tatango. Text Message Spam Infographic. 2011. Available online: https://www.tatango.com/blog/textmessagespam-infographic/ (accessed on 15 June 2020).
[4]. Goel, D.; Jain, A. Smishing-Classifier: A Novel Framework for Detection of Smishing Attack in Mobile Environment. In Proceedings of the Smart and Innovative Trends in Next Generation Computing Technologies (NGCT 2017), Dehradun, India, 30–31 October 2017; pp. 502–512.
[5]. Dahou, A.; Xiong, S.; Zhou, J.; Haddad, M.H.; Duan, P. Word Embeddings and Convolutional Neural Network for Arabic Sentiment Classification. In Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; The COLING 2016 Organizing Committee: Osaka, Japan, 2016; pp. 2418–2427.
[6]. Al-Smadi, M.; Qawasmeh, O.; AlAyyoub, M.; Jararweh, Y.; Gupta, B. Deep Recurrent neural network vs. support vector machine for aspect-based sentiment analysis of Arabic hotels’ reviews. J. Comput. Sci. 2018, 27, 386–393. [CrossRef] 19. Zhou, C.; Sun, C.; Liu, Z.; Lau, F.C.M. A-C-LSTM Neural Network for Text Classification. arXiv 2015, arXiv:cs.CL/1511.08630.