Report On

Email/SMS Spam Detection

Submitted in partial fulfillment of the requirements of the Course project in Semester VII of Final Year Computer Engineering

by

Dimple Khuman (Roll No. 04)
Divya Patil (Roll No. 08)
Aditi Sawant (Roll No. 12)

Mentor
Dr. Megha Trivedi

University of Mumbai

Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

(2023-24)

Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

CERTIFICATE

This is to certify that the project entitled “Email Spam Detection” is a bonafide work of "Dimple Khuman (Roll No. 04), Divya Patil (Roll No. 08), Aditi Sawant (Roll No. 12)" submitted to the University of Mumbai in partial fulfillment of the requirement for the Course project in semester VII of Final Year Computer Engineering.

Dr. Megha Trivedi
Mentor

Dr. Megha Trivedi	Dr. H.V. Vankudre
Head of Department	Principal

Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Course Project Approval

This Course Project entitled **“Email/SMS Spam Detection”** by **Dimple Khuman (Roll No. 04), Divya Patil (Roll No. 08), Aditi Sawant (Roll No. 12)** is approved for the degree of **Bachelor of Engineering** in Semester VII of Final Year **Computer Engineering .**

Examiners

1……………………………………..
(Internal Examiner Name & Sign)

2…………………………………………
(External Examiner Name & Sign)

Date:

Place:

Contents

Abstract

Acknowledgement

1 Introduction

1.1 Introduction

1.2 Problem Statement

1.3 Objectives

1.4 Scope

2 Literature Survey

2.1 Survey of Existing System

2.2 Limitation of Existing System

2.3 Mini Project Contribution

3 Proposed System

3.1 Introduction

3.2 Architecture

3.3 Algorithm and Process Design

3.4 Details of Hardware & Software

3.5 Experiment and Results for Validation and Verification

3.6 Conclusion and Future work.

3.7 References

Abstract

In an era dominated by the digital age, our communication landscape has undergone a remarkable transformation. Email and Short Message Service (SMS) have emerged as fundamental channels for personal and professional interactions. The ease and ubiquity of these platforms have, however, also made them fertile grounds for unwanted and often malicious messages – spam. Unwanted communications in the form of spam have been a persistent and growing concern for individuals and organizations alike. From product advertisements and phishing attempts to malicious software downloads and fraudulent schemes, spam poses significant risks to our digital lives. As the prevalence of spam grows, so too does the need for robust and efficient spam detection methods. In response to this challenge, the "Email/SMS Spam Detection" project endeavors to tackle this issue head-on, leveraging the power of Natural Language Processing (NLP) and machine learning.

Acknowledgement

We would like to extend our heartfelt gratitude to several individuals who played pivotal roles in the successful completion of our project. First and foremost, we express our deepest appreciation to Dr. Megha Trivedi, who served as both our Mentor and Head of Department. Her guidance, expertise, and unwavering support were instrumental throughout our research journey. We also acknowledge the invaluable contributions of our esteemed Principal, Dr. H.V. Vankudre, for their encouragement and belief in our capabilities. Furthermore, we cannot overlook the significant influence of our friends and family, who provided both direct and indirect assistance, serving as pillars of strength during challenging times. Additionally, we are indebted to the numerous academicians who generously served as supervisors and assessors, offering their expertise and insights that greatly enriched our project. Their collective efforts have been indispensable, and we express our sincere gratitude to each of them for making this study a resounding success.

1. Introduction

1.1 Introduction

The "Email/SMS Spam Detection" project addresses the problem of spam, which poses numerous challenges and risks to users and organizations. Spam messages encompass a wide spectrum of content, ranging from unsolicited product advertisements to phishing attempts, malware distribution, and fraudulent schemes. Spam messages often contain malware, phishing attempts, or fraudulent content that can harm individuals or organizations. Detecting and filtering out such messages is crucial to protect users from potential cyber threats. Spam messages can inundate email and messaging platforms, leading to wasted time and resources for individuals and businesses. By automating the identification and removal of spam, users can focus on legitimate messages. The volume of spam can overwhelm users, making it difficult to find and respond to important messages. Spam messages often aim to extract personal information or login credentials. Detecting and blocking these messages helps safeguard sensitive data and privacy. Spam can consume server space and network bandwidth, increasing infrastructure costs. Effective spam detection can reduce these operational expenses.

1.2 Problem Statement

The ubiquity of email and SMS as communication channels makes them highly attractive to spammers. According to the Radicati Group, over 319 billion emails were sent and received daily in 2021, and this number continues to rise. An alarming 45% of these emails are classified as spam. The problem is not limited to email; SMS spam is also on the rise, with spammers exploiting the simplicity and accessibility of text messaging to reach a broad audience. A significant portion of spam messages carries security threats. Malicious attachments, links to phishing websites, and fraudulent claims lure users into compromising their privacy and security. Opening a spam message can expose users to malware, ransomware, and identity theft. The consequences of falling victim to these security threats are dire, both for individuals and organizations. For businesses, allowing spam to reach customers' inboxes can have detrimental effects on brand reputation. Customers are more likely to lose trust in companies that fail to protect them from spam. This loss of trust can impact customer loyalty and ultimately revenue. The need for an effective and comprehensive spam detection system is evident, and this project seeks to provide a solution.

1.3 Objectives

1. Developing a robust and accurate spam detection system.
2. Creating a user-friendly interface for interaction.
3. Improving efficiency and automation in spam identification.
4. Enhancing cybersecurity by preventing malicious content.
5. Maintaining a positive brand image for organizations.
6. Enhancing user experience through spam-free communications.

1.4 Scope

1. Spam Detection: The project's primary scope is the development of a robust spam detection system that can accurately classify incoming messages as "Spam" or "Not Spam." This involves the use of machine learning and natural language processing techniques.
2. User-Friendly Interface: The system includes a user interface that is intuitive and user-friendly. Users can easily input messages for classification and receive classification results.
3. Cybersecurity Features: In addition to spam detection, the system enhances cybersecurity by identifying and preventing messages carrying malicious content, such as malware and phishing attempts.
4. Resource Optimization: The project focuses on optimizing server space and network bandwidth by efficiently filtering and managing spam. This can lead to cost savings and resource efficiency.
5. Machine Learning Model Refinement: Continuous refinement and improvement of the machine learning model are integral to the project's scope to ensure high accuracy and adaptability to changing spam tactics.
6. Future Expansion and Innovation: The project sets the stage for future expansion and innovation in the field of spam detection. Future work may include advanced machine learning techniques, multilingual support, integration with messaging apps, and more.

2 Literature Survey

2.1 Survey of Existing System

1. Lack of Effective Spam Detection: Many email and SMS platforms currently lack robust spam detection mechanisms, leading to users being inundated with unwanted messages.
2. Resource Drain: Existing systems often struggle with resource inefficiencies, as they are overwhelmed by the volume of spam messages, resulting in wasted server space and network bandwidth.
3. User Frustration: Users are frustrated by the continuous influx of spam, which distracts them from legitimate communications and compromises their privacy and security.

2.2 Limitation of Existing System

1. Rule-Based Filters with Limited Adaptability: Many existing systems rely on rule-based filters that use predefined patterns to identify spam. These rules might not adapt well to evolving spam tactics and patterns, leading to false positives or false negatives.
2. Lack of Real-Time Processing: Some existing systems lack real-time processing capabilities, which means that users might not receive immediate feedback on the classification of messages.
3. Resource Inefficiencies: Existing systems may suffer from resource inefficiencies due to the sheer volume of spam messages, leading to wasted server space, network bandwidth, and computational power.
4. Privacy and Security Concerns: Users might express concerns about privacy invasion and the presence of potentially harmful content in their inboxes, as existing systems struggle to protect them adequately.
5. Need for Continuous Model Improvement: Existing machine learning models used for spam detection may not always adapt to emerging spam techniques. They require continuous updates and refinements.
6. Addressing Challenges Specific to SMS Spam: While there's substantial research on email spam detection, SMS spam poses unique challenges, including character limitations and limited metadata. Existing solutions may not adequately address these specific issues.

2.3 Mini Project Contribution

The project can contribute to a more seamless and efficient user experience by automating the identification and management of spam. Users can engage with their email and SMS messages without the distraction of unwanted content. By detecting and blocking spam messages that carry malicious content, the project can significantly enhance cybersecurity. This, in turn, protects users from malware, phishing attempts, and other security threats. Detecting and blocking messages that aim to collect personal information can contribute to user privacy and data protection. Users can trust that their sensitive information remains secure. The project can provide educational resources and tutorials to help users understand the system's features and best practices for spam management. This contributes to user awareness and digital literacy. The project can be designed to be scalable, capable of handling a variety of user volumes and message types. This scalability is an essential feature for addressing the dynamic nature of spam.

3 Proposed System

3.1 Introduction

In an age where digital communication is the lifeblood of our personal and professional interactions, the relentless invasion of unwanted messages has become an unwelcome reality. Whether through email or Short Message Service (SMS), the digital world has given birth to spam—a deluge of irrelevant, distracting, and potentially dangerous messages. To address this growing challenge and elevate the digital communication experience, we introduce the "Email/SMS Spam Detection" system. The digital landscape has evolved at an astounding pace, making email and SMS fundamental channels for communication. These platforms, lauded for their accessibility and immediacy, connect us across continents and serve as essential tools in both our personal and professional lives. Yet, the very attributes that make them indispensable have rendered them vulnerable. The consequence: a barrage of spam messages that inundate our inboxes and disrupt the fluidity of our interactions. The "Email/SMS Spam Detection" system is our response to this pressing issue. It is a step towards ensuring the sanctity and security of our digital communications. This proposed system addresses the critical challenges posed by spam and endeavors to enhance the digital communication experience in multiple dimensions.

3.2 Architecture

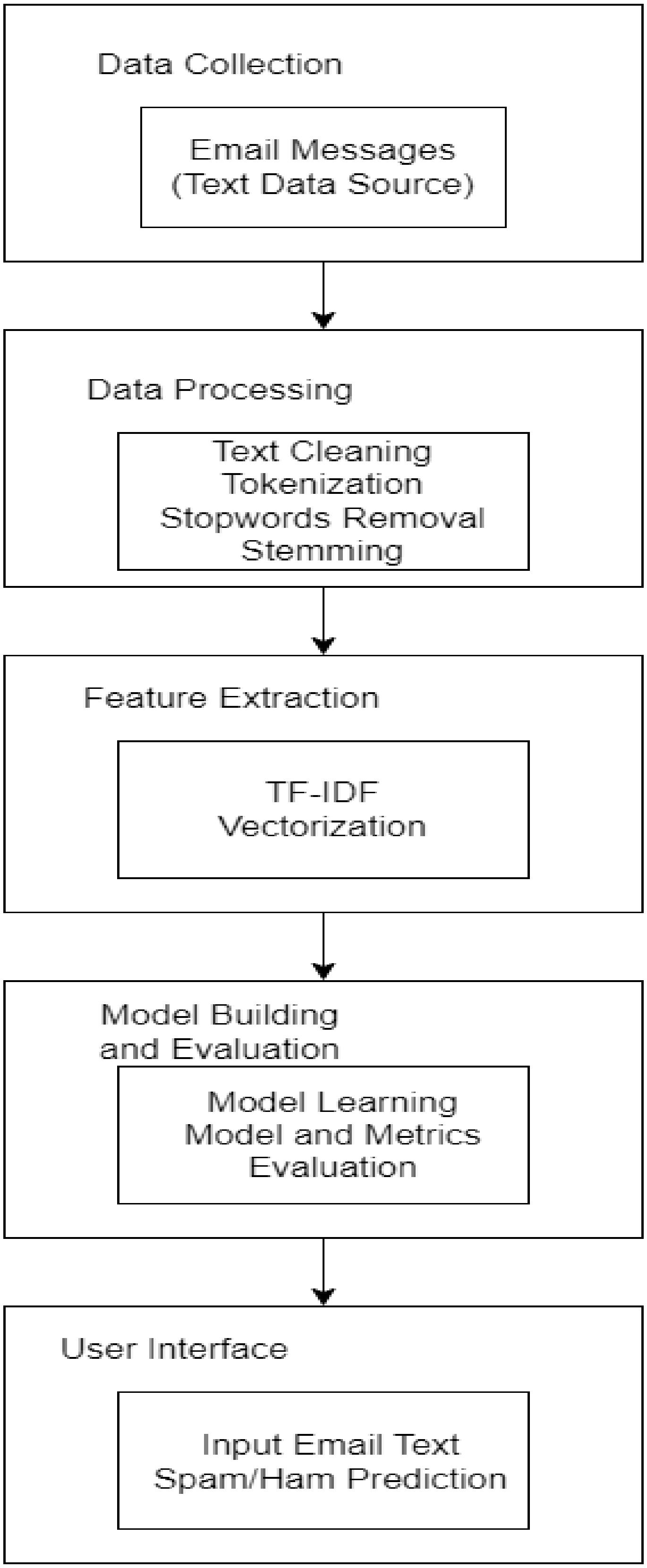


Fig. 3.2 Spam Detection Architecture

3.3 Algorithm and Process Design

1. Data Collection:
    a. Collect email messages from various sources, including user input and email datasets.
    b. Store emails in a structured format, such as a data frame or a database.
2. Data Preprocessing:
    a. For each email message, perform the following preprocessing steps:
        i. Remove special characters, HTML tags, and other irrelevant content.
        ii. Convert the email text to lowercase to ensure uniformity.
        iii. Tokenize the text by splitting it into individual words.
        iv. Remove common stopwords like "the," "and," "is," etc.
        v. Apply stemming or lemmatization to reduce words to their root form.
3. Feature Extraction:
    a. Convert preprocessed email text into numerical feature vectors using the TF-IDF (Term Frequency-Inverse Document Frequency) method:
        i. Calculate the TF-IDF scores for each word in the email.
        ii. Build a TF-IDF matrix, where rows represent emails and columns represent unique words.
4. Model Selection:
    a. Choose an appropriate machine learning or natural language processing model for classification.
    b. Split the data into training and testing sets for model evaluation.
5. Model Training:
    a. Train the selected model using the training data.
    b. The model learns to classify emails into spam or ham based on the features derived from TF-IDF.
6. Model Evaluation:
    a. Evaluate the model's performance on the testing data using metrics like accuracy, precision, recall, F1-score, and a confusion matrix.
7. User Interface:
    a. Create a user-friendly interface where users can input email text.

    b. Users can trigger the spam classification process by submitting an email through this interface.

8. Spam Classification:
    a. When an email is submitted through the user interface, the system performs the following steps:
        i. Preprocess the email text.
        ii. Extract features using the pre-trained TF-IDF vectorizer.
        iii. Use the trained model to predict whether the email is spam or ham.
        iv. Present the prediction to the user (spam or ham).

3.4 Details of Hardware & Software

Hardware Specifications:

- System type: x64-based processor, 64-bit operating system.
- Memory (RAM) installed: 8.00 GB (7.34 GB Usable)
- Total size of Hard disk: 1 TB

Software Specifications:

- Operating system: Microsoft Windows 10
- Integrated Development Environment: PyCharm
- Streamlit, Scikit-Learn, NLTK(Natural Language Toolkit)
- Programming language: Python

3.5 Experiment and Results for Validation and Verification:

Exploratory Data Analysis **:**

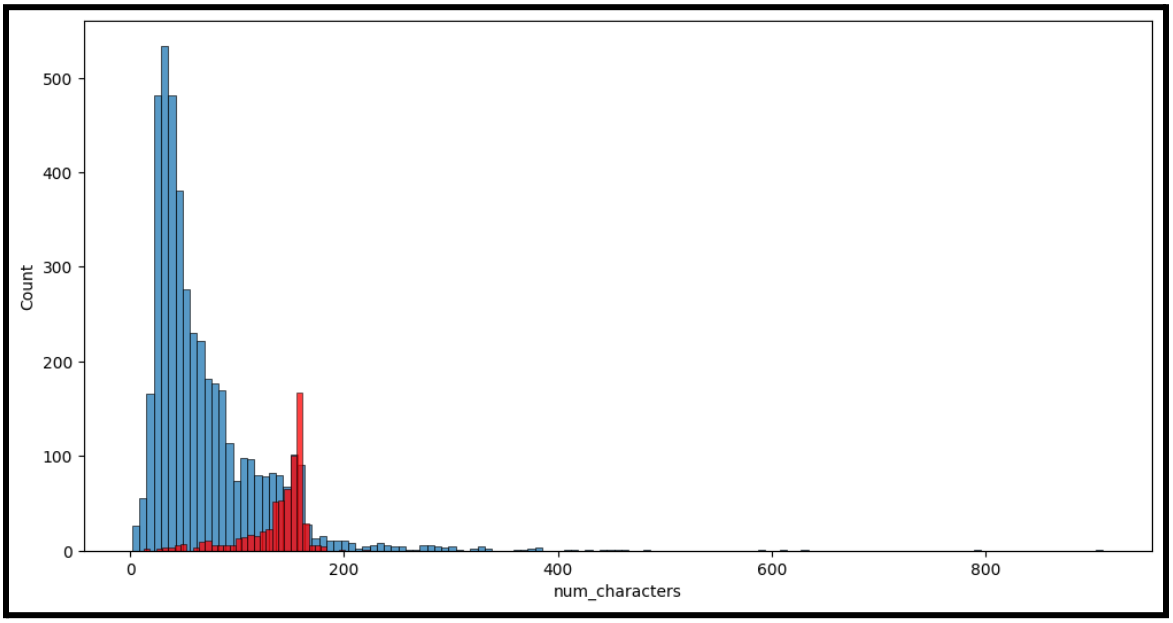


Fig. 3.5.1 Histogram for Character Comparison for ham and spam

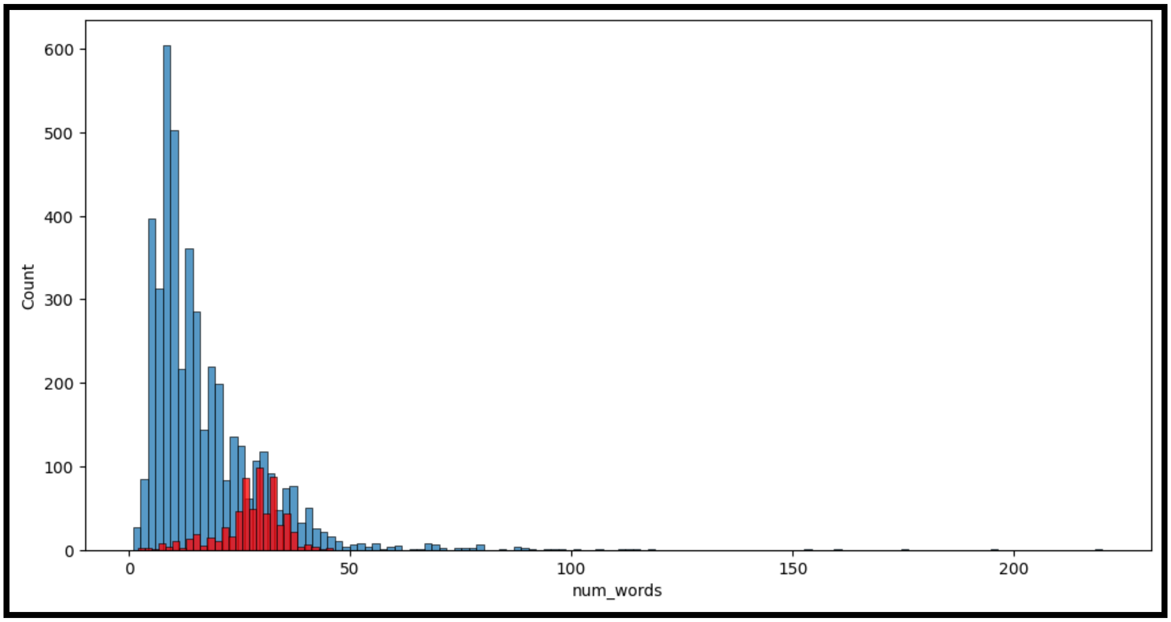


Fig. 3.5.2 Histogram for Words comparison for ham and spam

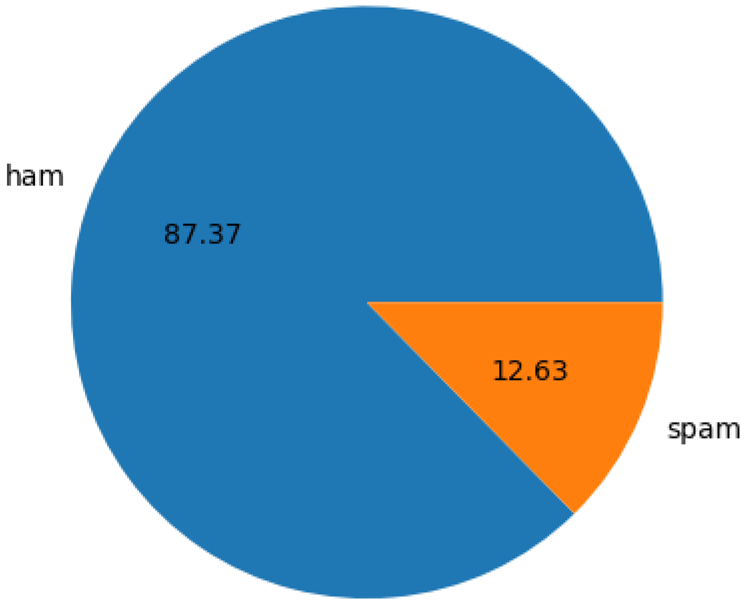
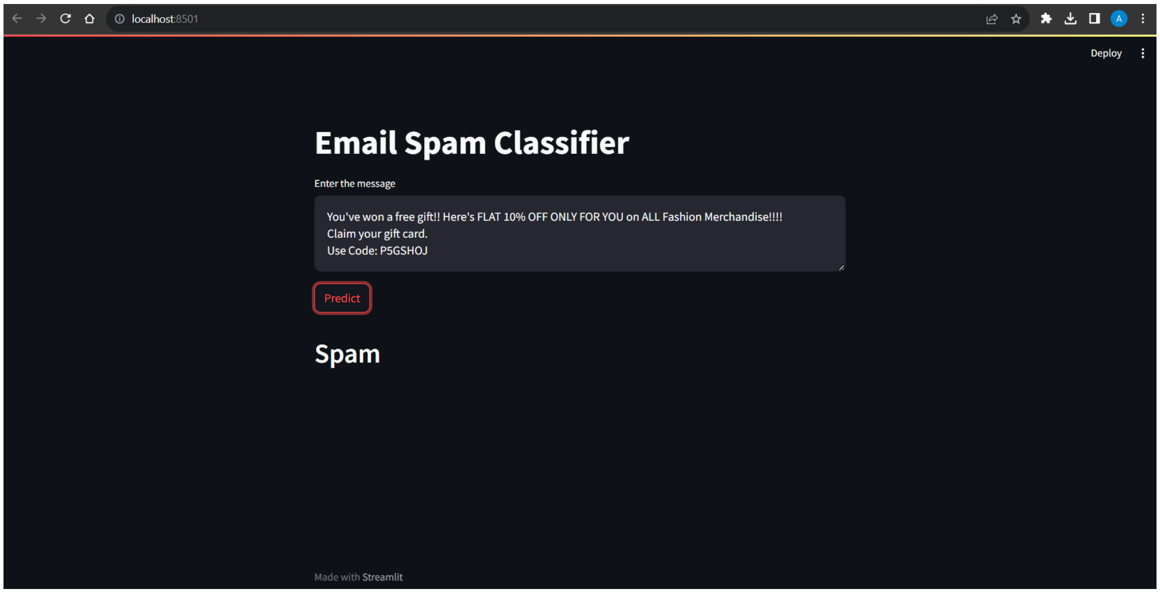


Fig. 3.5.3 Pie chart for Ham and Spam



Email Spam Classifier

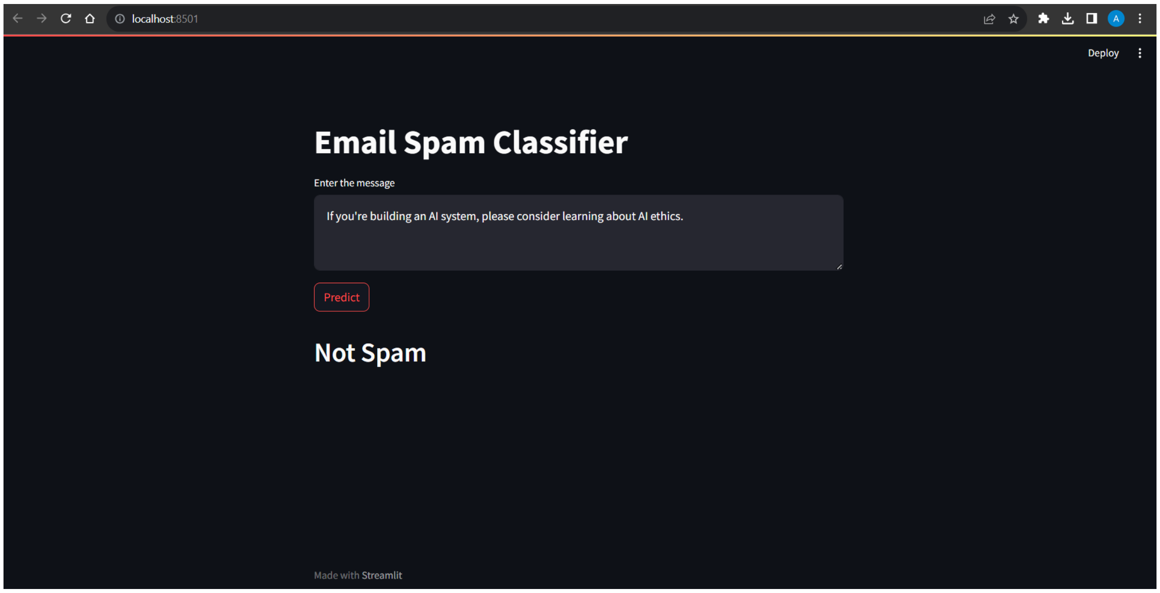
Enter the message

You've won a free gift!! Here's FLAT 10% OFF ONLY FOR YOU on ALL Fashion Merchandise!!!!
Claim your gift card.
Use Code: P5GSHOJ

Spam

Fig. 3.5.4 Output showing as Spam Email

When a message or email is classified as spam by a spam filter or spam detection system, the output typically involves marking or flagging the message as spam.



Email Spam Classifier

Enter the message

If you're building an AI system, please consider learning about AI ethics.

Not Spam

Fig. 3.5.5 Output showing as Not a Spam Mail

When an email or message is not classified as spam by your NLP and machine learning software, the output or result will typically indicate that the message is non-spam or ham.

3.6 Conclusion and Future work.

The proposed "Email/SMS Spam Detection" system represents a significant leap forward in the realm of digital communication. This system aims to address the ever-growing challenge of spam, which disrupts our inboxes, poses security risks, and compromises the efficiency of digital interactions. By implementing advanced machine learning algorithms, user-friendly interfaces, and customization options, the system offers a comprehensive solution to these issues. With the ability to identify and classify messages as "Spam" or "Not Spam," the system empowers users to regain control over their digital communication environment. By providing real-time processing, privacy protection, and enhanced cybersecurity features, it offers a robust defense against malicious content.

Future work:

1. Advanced Machine Learning: Continuously improve the system's spam detection model by exploring advanced machine learning techniques, such as deep learning, reinforcement learning, and transfer learning.
2. Multilingual Support: Extend the system's capabilities to handle multilingual content effectively, recognizing that spam is a global issue.
3. Enhanced User Experience: Further refine the user interface and customization options to provide an even more intuitive and personalized user experience.
4. Real-Time Updates: Implement mechanisms for real-time updates of the spam detection model to adapt to rapidly evolving spam tactics.
5. Integration with Messaging Apps: Expand the system's reach by integrating it with popular messaging apps and platforms to address spam in various channels.

3.7 References

[1] Kingshuk Debnath and Nirmalya Kar. Email Spam Detection using Deep Learning Approach. 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON).

[2] Aakash Atul Alurkar, Sourabh Bharat Ranade, Shreeya Vijay Joshi, Siddhesh Sanjay Ranade, Piyush A. Sonewar, Parikshit N. Mahalle and Arvind V. Deshpande. A proposed data science approach for email spam classification using machine learning techniques. 2017 Internet of Things Business Models, Users, and Networks.

[3] Mansoor RAZA, Nathali Dilshani Jayasinghe and Muhana Magboul Ali Muslam. A Comprehensive Review on Email Spam Classification using Machine Learning Algorithms. 2021 International Conference on Information Networking (ICOIN).

[4] Prazwal Thakur, Kartik Joshi, Prateek Thakral and Shruti Jain. Detection of Email Spam using Machine Learning Algorithms: A Comparative Study. 2022 8th International Conference on Signal Processing and Communication (ICSC).

[5] Sanaa Kaddoura, Omar Alfandi and Nadia Dahmani. A Spam Email Detection Mechanism for English Language Text Emails Using Deep Learning Approach. 2020 IEEE 29th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE).