iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62390

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 V May 2024- Available at www.ijraset.com

Machine Learning Based Email Spam Detection: Achieving High Accuracy and Efficiency

Pushkar Joglekar[1], Janhavi Rajurkar[2], Madhuri Shinde[3], Pranav Tayde[4], Ved Gadmade[5]

Abstract: ***Email communication has become an essential aspect of modern-day interactions, but the proliferation of spam emails poses significant challenges to users' productivity and security. This research paper presents a comprehensive study on the development and implementation of an efficient email spam detection and categorization system. The project aims to categorize emails into predefined sections by using the Support Vector Machine (SVM) model, Flask, and the Gmail API, ensuring accuracy and efficiency in email classification. The methodology involves data preparation, processing, storage, and management, ensuring robust security and privacy considerations. The system's three-tiered classification strategy enhances the accuracy of spam and ham detection. Future enhancements include integrating advanced machine learning models, user feedback mechanisms, and multi-platform support to adapt to evolving email trends and user preferences. This research contributes to the field of email management by offering a new approach to combat spam effectively and enhance email organization for users in the digital age.***

Keywords: Machine Learning, Email Classification, Spam Detection, Flask, Gmail API, Support Vector Machine, Text Processing.

I. INTRODUCTION

Email is indispensable for modern communication but is often disrupted by spam, which clutters inboxes and threatens productivity and security. This paper presents an innovative email spam detection and categorization system using the Support Vector Machine (SVM) model. By integrating SVM with the Flask web framework and Gmail API, the system aims to enhance the accuracy and efficiency of email classification, improving user experience and inbox organization.

The paper unequivocally proposes the following:

1) Developing an advanced email spam detection and categorization system using machine learning techniques.
2) Enhancing user inbox organization and email management by effectively filtering out spam emails and categorizing legitimate emails based on content analysis, sender information, and subject analysis.
3) Ensuring robust security and privacy considerations in data processing, storage, and management while incorporating advanced machine learning models, user feedback mechanisms, and multi-platform support to adapt to evolving email trends and user preferences.

By addressing the limitations of traditional spam detection methods, such as rule-based filtering, this research aims to provide a dynamic and effective approach to classifying emails. The methodology employed encompasses a multi-faceted approach to email management, including data preparation, processing, and storage, with a strong emphasis on security and privacy considerations. The system's classification strategy, analyzing Email ID, Subject, and Message Content, enhances the precision of spam detection and categorization. Furthermore, its ability to categorize emails into distinct groups offers users a streamlined approach to organizing their inboxes and prioritizing essential communications. Through the integration of advanced machine learning models and user feedback mechanisms, the system adapts to evolving email trends and user preferences, ensuring its relevance and effectiveness in the ever-changing landscape of digital communication.

II. LITERATURE REVIEW

Through an extensive review of existing literature, this research explores various approaches to implementing effective encryption and decryption algorithms for enhancing security. By examining key findings from several studies, this paper aims to identify the most suitable methodologies for ensuring robust data protection. The paper [1] highlights the importance of spam detection using machine learning, particularly the potential of hybrid algorithms and deep learning for improved accuracy, and suggests incorporating blockchain and linguistic expertise for better dataset annotation. In their study [2], researchers propose innovative algorithms such as WITCH, semi-supervised learning link-based algorithms, and K-Nearest Neighbor for feature extraction to improve accuracy and efficiency in email spam detection.

The paper [3] investigates spam email detection in Urdu using machine learning algorithms like Naive Bayes, CNN, SVM, and LSTM, with the LSTM model achieving 98.4% accuracy, addressing challenges due to the unique script and dataset scarcity, and contributes by creating a dedicated Urdu dataset for ML and DL models. The paper [4] discusses email spam filtering techniques, emphasizing machine learning methods like Naive Bayes, SVM, K-Nearest Neighbor, and evaluates different systems to highlight the importance of effective filtering and the need for advanced techniques to combat spam. The paper [5] explores spam email detection using machine learning algorithms optimized with bio-inspired methods, finding Multinomial Naïve Bayes with Genetic Algorithm to perform best, and compares various techniques to determine the most suitable model.

The paper [6] proposes a combined approach using the PV-DM Neural Network model to enhance spam filters by capturing context and relevant features, outperforming traditional BOW and PV-DM methods on Enron and Ling spam datasets. The paper [7] highlights the use of NLP and machine learning, particularly the BERT transformer model, for spam email detection, demonstrating its superior accuracy and F1 score compared to classic classifiers and baseline DNN models, emphasizing the importance of these technologies in cybersecurity.

The paper [8] addresses the pressing issue of spam email, which accounts for a significant portion of global email traffic, causing inconvenience and financial losses to users. It emphasizes the importance of automatic email filtering and highlights the escalating battle between spammers and filtering methods. Two main approaches, knowledge engineering, and machine learning, are discussed, with machine learning being favored for its efficiency and ability to adapt without the need for constant rule updates. The study reviews several machine learning algorithms, including Naïve Bayes, support vector machines, Neural Networks, K-nearest neighbor, Rough sets, and the artificial immune system, assessing their applicability to spam email classification.

The paper [9] explores SMS spam detection using a bagging approach with RVM, SVM, Naive Bayes, and KNN algorithms, highlighting the effectiveness of the RVM model, which achieved the best performance with an F1 score of 0.975175, and emphasizing the importance of accurate SMS spam classification. The literature review of the paper [10] focuses on the increase in email usage for business transactions and communication, highlighting the vulnerability of emails to spam attacks. It discusses the importance of data mining in addressing the spam issue and the use of classification algorithms for spam detection. The review also emphasizes the need for accurate classification algorithms to improve email filtering efficiency.

III. METHODOLOGY

The methodology employed in our email spam detection system is a comprehensive approach that blends software engineering principles with machine learning techniques and data analysis. This project aims to develop a robust system capable of effectively distinguishing between spam and legitimate emails. The methodology is structured into several key components, including setting up a web application using Flask, integrating the Gmail API for email access, processing and preparing email data, and employing a Support Vector Machine (SVM) for classification tasks.

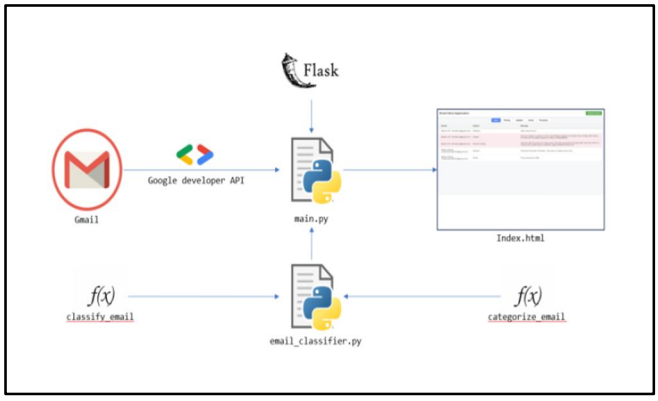


Fig. 1 System Architecture of the proposed model

The architecture of the system depicted in Fig. 1 consists of three main components: Gmail API for email retrieval, Python scripts for email classification, and a Flask web application for displaying the classification results. The proposed system is an email classification system designed to classify emails as spam or non-spam (ham). The system utilizes the Gmail API to fetch emails from a user's Gmail account and processes them using Python scripts, including a Flask-based web application. The workflow begins with the Gmail API fetching emails, which are then passed to Python scripts for classification.

The classification results are then displayed on a web page using HTML. The system likely uses machine learning algorithms, such as Support Vector Machines (SVM), for email classification. After establishing the system architecture, the development of the email spam detection system proceeded with a focus on each component's implementation and integration. The following sections outline the development process for the web application, email access and processing, data preparation, and machine learning model integration.

A. Web Application Development

The front end of the email spam detection system was developed using HTML, CSS, and minimal JavaScript to create a user-friendly interface for displaying email classifications. The development process involved designing the layout and visual elements of the interface to ensure clarity and ease of use. Flask's template rendering capabilities were utilized to integrate the front end with the back end, allowing for the dynamic display of email data fetched from the Gmail API. Overall, the front end was designed to provide users with an organized and intuitive interface for interacting with the system.

The backend of the system was developed using Python with Flask to handle HTTP requests, manage routes, and process email data. Flask's lightweight and flexible nature made it ideal for building the backend of the web application. The backend was responsible for integrating with the Gmail API to fetch emails, processing and preparing email data for classification and integrating with the SVM model for email classification. Additionally, the backend managed the communication between the front end and the machine learning model, ensuring that the classified email data was displayed correctly on the front end. Overall, the backend was crucial for the functionality of the email spam detection system, handling the core processing and logic of the application.

B. Email Accessing and Processing

The email access and processing component of the system involves integrating the Gmail API to access user emails and parsing these emails to extract crucial elements. The first step in this process is setting up OAuth 2.0 credentials, which involves creating a project in the Google Cloud Console, enabling the Gmail API for the project, and generating client credentials. These credentials are then used by the Flask web application to authenticate with the Gmail API. Once authenticated, the Flask app can fetch emails from the user's Gmail account using the Gmail API's messages. list` and `messages. get` methods.

Email parsing is the next step, where the fetched emails are processed to extract key information such as sender, subject, and message body. This is achieved by iterating through the email messages and extracting relevant information using Python's email module. Different email formats and content types, such as plain text, HTML, and attachments, are handled using appropriate parsing techniques. For example, email bodies in HTML format are parsed to extract text content, while attachments are processed separately to extract their contents or metadata. Overall, the integration of the Gmail API and email parsing ensures that the system can access and process user emails effectively for spam detection.

C. Data Preparation

In the data preparation phase, a labeled dataset of emails is utilized for training and testing the Support Vector Machine (SVM) model, which is crucial for the machine learning process. This dataset consists of emails that are tagged as either spam or ham (non-spam). The importance of this dataset lies in its role as the foundation for the SVM model's training, allowing it to learn patterns and characteristics of spam and non-spam emails.

Preprocessing techniques are applied to convert textual data into a suitable format for machine learning algorithms, involving steps like removing stop words, tokenizing, and stemming or lemmatization to ensure accuracy and prepare the data for analysis. Techniques such as tokenization, which involves splitting the text into individual words or tokens, and TF-IDF (Term Frequency-Inverse Document Frequency) vectorization, which assigns weights to words based on their frequency and importance in the dataset, are used. These techniques help transform the textual data into numerical features that can be used by the SVM model for classification. Overall, the data preparation phase ensures that the SVM model is trained on a well-prepared dataset, enhancing its ability to accurately classify emails as spam or ham.

D. Machine Learning Model

In the machine learning model component, the Support Vector Machine (SVM) was chosen for its effectiveness in high-dimensional spaces and suitability for text classification problems. The SVM model was trained on a portion of the dataset and tested on a separate set to evaluate its accuracy, precision, and recall.

Feature engineering techniques were applied to enhance the model's ability to distinguish between spam and ham emails accurately. Overall, the SVM model selection, training process, and feature engineering efforts were crucial in developing a robust email spam detection system.

	Accuracy	Precision	Recall	F1-Score
Multinomial Naive Bayes	0.952519	0.954981	0.952519	0.947935
Logistic Regression	0.970930	0.971237	0.970930	0.969602
Support Vector Machine	0.985465	0.985447	0.985465	0.985203
Random Forest	0.978682	0.978768	0.978682	0.978033
K-Nearest Neighbors	0.907946	0.916770	0.907946	0.885559

Fig. 2 Algorithm Performance Evaluation over Multiple Metrics

The observed results shown in Fig. 2 from the comparison of machine learning algorithms show that the Support Vector Machine (SVM) outperformed other algorithms, including Multinomial Naive Bayes, Logistic Regression, Random Forest, and K-Nearest Neighbors, in terms of Accuracy, Precision, Recall, and F1-Score. With an Accuracy of 0.985465, Precision of 0.985447, Recall of 0.985465, and F1-Score of 0.985203, SVM demonstrated superior performance in this classification task.

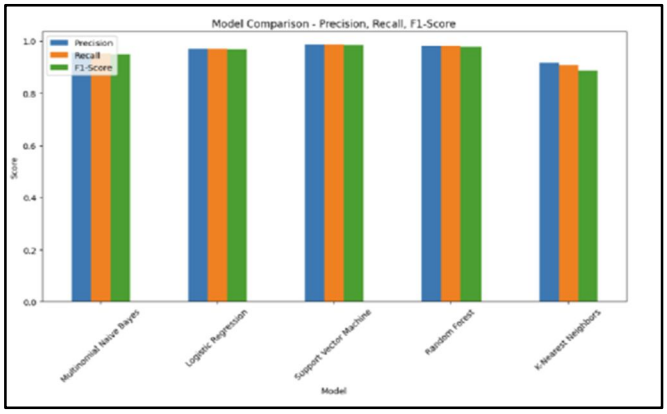


Fig. 3 Comparative Performance Metrics of Different Models

The bar graph in Fig. 3 indicates that the Support Vector Machine (SVM) is the top-performing model across precision, recall, and F1-score metrics, consistently outperforming other models. Logistic Regression also performs well, particularly in precision and F1-score, while Naive-Bayes Bernoulli and Random Forest exhibit similar performance levels. Neural Network, although respectable, falls behind the top-performing models.

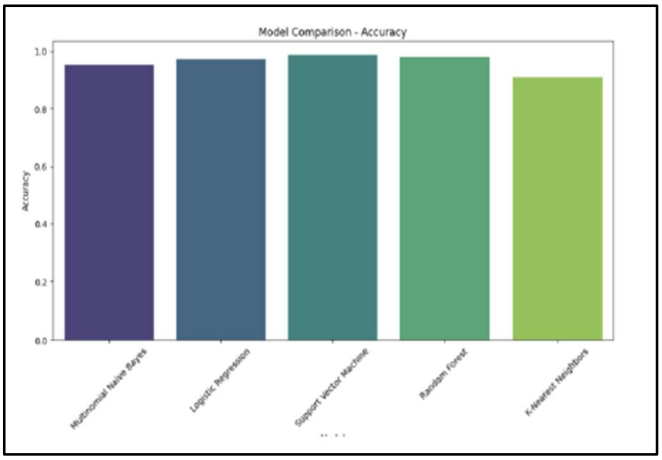


Fig. 4 Model Comparison - Accuracy

The bar graphene Fig. 4 illustrates the accuracy comparison of five models, with "Neural Network with BN" achieving the highest accuracy, followed by "Random Forest," "Logistic Regression," and "Support Vector Machine," while "k-Nearest Neighbors" lags with the lowest accuracy. These results suggest that "Neural Network with BN" performs the best among the models evaluated in this comparison.

Thus, Support Vector Machines (SVMs) are often considered the best choice for text classification tasks like email spam detection due to their ability to handle high-dimensional data efficiently. SVMs excel in scenarios where the number of dimensions exceeds the number of samples, which is common in text analysis. Their effectiveness lies in their capacity to find the optimal hyperplane that maximizes the margin between classes, leading to better generalization and performance on unseen data. SVMs have shown exceptional performance in accurately differentiating between spam and legitimate emails, as evidenced by their high accuracy and F1-score in such tasks. Overall, SVMs are a reliable and effective choice for text classification problems, making them the preferred model for many applications.

IV. RESULTS AND DISCUSSIONS

The results of our email spam detection project showcase the effectiveness and robustness of our system in accurately classifying and categorizing emails. Through a combination of rule-based filtering, machine learning techniques, and natural language processing methods, our system has demonstrated a high level of accuracy in distinguishing between spam and legitimate emails. These results underscore the potential of our approach to address the persistent challenge of email spam, offering a reliable solution for users seeking to manage their email communication more efficiently.

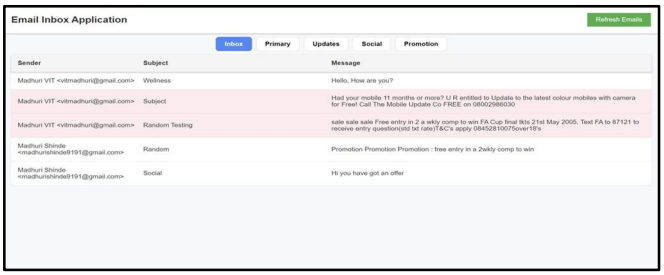
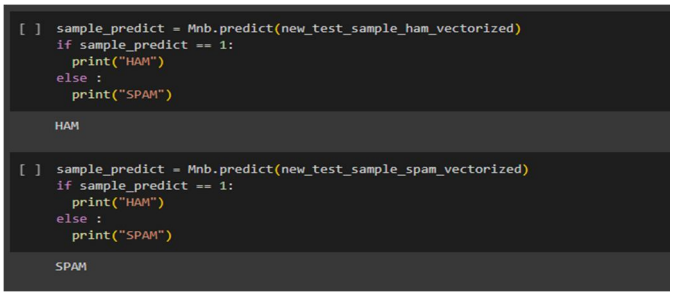


Fig. 5 GUI Interface

In the results of our research, the email classification system effectively categorizes emails into two main groups: spam and ham (non-spam). Spam emails are identified by their characteristic content, such as offers for free products or services, urgency in prompts for action, requests for personal information, and the presence of keywords commonly associated with spam.

On the other hand, ham emails exhibit more genuine and non-spam characteristics, such as personal greetings, the absence of aggressive promotional language, and contextual messages that suggest genuine conversation or social interaction. The classification criteria likely include content analysis, keyword detection, and possibly machine learning algorithms to identify patterns typical of spam emails. Overall, this system plays a crucial role in email management by ensuring that users are protected from unsolicited or harmful content while allowing legitimate emails to reach their intended recipients.



```
[ ]  sample_predict = Mnb.predict(new_test_sample_ham_vectorized)
     if sample_predict == 1:
       print("HAM")
     else :
       print("SPAM")

     HAM

[ ]  sample_predict = Mnb.predict(new_test_sample_spam_vectorized)
     if sample_predict == 1:
       print("HAM")
     else :
       print("SPAM")

     SPAM
```

Fig. 6 Metrics Demonstration of model for identifying Spam or Ham.

The Naive Bayes classifier shown in Fig.6 performed well on the test samples, correctly identifying a non-spam (HAM) email as "HAM" and a spam email as "SPAM." This suggests that the model can effectively differentiate between spam and non-spam emails based on the features it has learned. The accurate classification of both types of emails indicates the model's potential reliability in real-world email classification tasks.

V. FUTURE SCOPE

Our project aims to stay ahead in the dynamic email landscape by continuously refining our machine-learning models. We plan to incorporate advanced techniques beyond simple keyword analysis to improve the accuracy of our email classification system. Additionally, we aim to integrate user feedback mechanisms to fine-tune our algorithms based on real-world interactions. This iterative process will ensure our system remains responsive to evolving user needs and preferences.

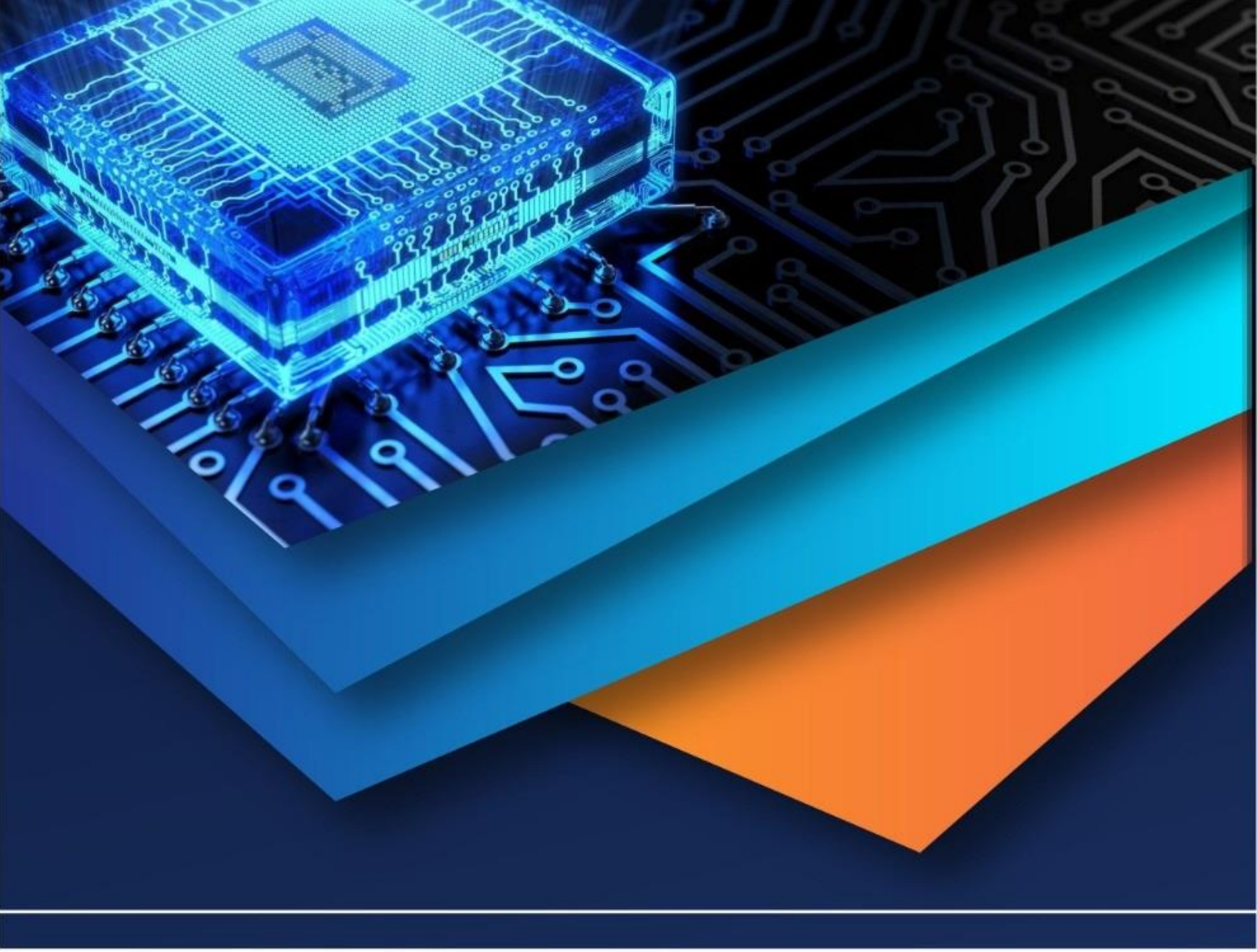
Expanding compatibility to support multiple email platforms is another key goal. This will enhance the accessibility of our solution, catering to a broader audience across various providers. User experience remains a top priority, and we are committed to maintaining intuitive, user-friendly interfaces. Security and privacy are paramount, and we will uphold the highest standards to safeguard user data against potential threats, building and maintaining trust with our users.

VI. CONCLUSION

In conclusion, our research has contributed a novel approach to email management by developing an efficient email spam detection and categorization system. Through a comprehensive review of existing methodologies and technological advancements, we identified the need for a more adaptable and accurate solution to combat evolving spam tactics. Our system, detailed in the "Project Development" section, leverages machine learning models and keyword-based checks to classify incoming emails as spam or legitimate. By incorporating features such as sender information, email content, and subject analysis, our system achieves a high level of accuracy. Additionally, we extended our research to categorize legitimate emails into primary, social, promotion, and update categories based on content analysis, enhancing user inbox organization and overall email management. Our work represents a significant advancement in the field of email management and provides a valuable contribution to combating the ever-evolving challenges of email spam.

REFERENCES

[1] N. Ahmed, "Machine Learning Techniques for Spam Detection in Email and IOT Platforms: Analysis and Research Challenges," Hindawi Security and Communication Networks, vol. 2022, 2022.
[2] R. Sharma, "E-Mail Spam Detection Using SVM and RBF," I.J. Modern Education and Computer Science, 2016,
[3] Z. B. Siddique, "Machine Learning-Based Detection of Spam Emails," Hindawi, vol. 2021, 2021.
[4] H. Bhuiyan, "A Survey of Existing E-Mail Spam Filtering Methods Considering Machine Learning Techniques," Global Journal of Computer Science and Technology: C Software & Data Engineering, vol. 18, no. 2, 2018.
[5] S. GIBSON, "Detecting Spam Email With Machine Learning Optimized With Bio-Inspired Metaheuristic Algorithms.," IEEE Access, vol. 8, 2020.
[6] S. Douzi, "Hybrid Email Spam Detection Model Using Artificial Intelligence," International Journal of Machine Learning and Computing, vol. 10, no. 2, 2020.
[7] I. AbdulNabi, "Spam Email Detection Using Deep Learning Techniques," in The 2nd International Workshop on Data-Driven Security (DDSW 2021) March 23 - 26, 2021, Warsaw, Poland, 2021.
[8] W. Awad, "MACHINE LEARNING METHODS FOR SPAM E-MAIL CLASSIFICATION," International Journal of Computer Science & Information Technology (IJCSIT), vol. 3, no. 1, 2011.
[9] S. Pudasainia, "SMS Spam Detection using Relevance Vector Machine," in 3rd International Conference on Evolutionary Computing and Mobile Sustainable Networks (ICECMSN 2023), 2023.
[10] S. M. Abdulhamid, "Comparative Analysis of Classification Algorithms for Email Spam Detection.," I. J. Computer Network and Information Security, vol. 1, 2018.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY