Smart Mail Mentor

*

Purushotham Kalangi
Computer Science
Seidenberg School Of CS and IS
Pace University
pk97681n@pace.edu

Sathvik Chowdary Samineni
Computer Science
Seidenberg School Of CS and IS
Pace University
ss63111n@pace.edu

Rakesh Kuchipudi
Computer Science
Seidenberg School Of CS and IS
Pace University
rk24434n@pace.edu

Kaleemunnisa, FNU
Computer Science
Seidenberg School Of CS and IS
Pace University
klnu@pace.edu

***Abstract*—In today's digital world, thousands of emails are sent and received every day, overloading people and increasing the danger of missing important communications. Traditional email management software often fails to adapt to user-specific needs, making manual organizing time-consuming and inefficient. To solve this issue, we present an AI-powered email summary and organizing system that improves email management via automation and customization. By leveraging Natural Language Processing (NLP) techniques to summarize emails within predefined time periods (daily, weekly, or monthly), extract critical information like deadlines, events, and meetings, and automatically incorporate them. Furthermore, machine learning algorithms leverage user preferences and previous behavior to categorize and prioritize emails, ensuring that critical communications are easily accessible. Initial testing shows that our AI-powered solution post classifying (TF-IDF) the emails with the combination of extractive and abstractive summarization techniques dramatically decreases the time spent on email management while enhancing the accuracy of vital information detection. Users report better efficiency and less inbox clutter, resulting in more productivity. This solution highlights the AI's potential to change digital communication by decreasing information overload and streamlining productivity. Future enhancements will center on improving NLP models for better contextual comprehension and expanding integrations with third-party productivity tools.**

***Index Terms*—Text classification, hybrid text summarization, TF-IDF, Random forest classifier, SMOTE, LSTM**

I. Introduction

In the early 1970s, email was invented, and from then email became mainstream for communication by enabling fast,global, and efficient messaging transforming business and personal interactions.Today, email is the main method we use for long-distance communication. More than 100 billion messages are received and sent every day, and that number is only expected to increase. As billions of messages are generated daily, the chance of missing messages by individual or organizations are at high risk.Missing the mails lead to a big loss for business or organizations or individuals.

Preventing the loss of messages or mails is crucial in the current world. Existing emails have lack of features that automate the generation of the mail report for daily or weekly. Smart mail mentor feature helps prevent the loss of messages.

Smart mail mentor is Natural language processing technique, artificial intelligence, machine learning, and deep learning. This feature automates daily or weekly emails and generates the report without losing mails.

II. Literature Review

Email management has become an essential task in the digital age of today, where individuals and organizations receive a large volume of emails daily. Efficiently managing and prioritizing emails is crucial to improving productivity and reducing information overload. Several studies have explored various approaches to automate email classification, prioritization, and summarization.

Machine learning (ML) techniques have played a significant role in email classification. Traditional approaches such as Naïve Bayes, Support Vector Machines (SVM), and Random Forest classifiers have been widely used to categorize emails into predefined categories such as spam, promotional, or business emails [1]. However, recent advances in deep learning, particularly long-short-term memory (LSTM) networks, have demonstrated improved accuracy in text-based classification tasks, including email filtering [2].

The TF-IDF (Term Frequency-Inverse Document Frequency) and Random Forest approach has been particularly effective in feature extraction and classification. TF-IDF helps identify important words in emails, while Random Forest improves classification accuracy by leveraging an ensemble of decision trees [3]. Furthermore, LSTM-based models have shown superior performance in handling sequential text data, making them well suited for email classification [4].

One of the areas to focus on to summarize and classify emails was explored in the paper [5] which exposes abstractive summarization using BERT-based extractive text summarization and BERT -GPT -2 hybrid models. We effectively managed to gather topic-aware summarization(T-BERTSum) that can potentially identify relevant summaries, enhancing organization and prioritization. Moreover, [5] suggests that by fine-tuning the BERT and GPT models in the paper, results could be obtained for the Q / A capabilities of our tool.

Interpreting the text and generating the human-like response by understanding the context and meaning by [6] tries to generate response by emphasizing both the abstractive and extractive approaches for more accurate and coherent summaries, this is helpful in giving the relevant information by using the technique that is seq-to-seq modelling using LSTM and encoder-decoder and however, [6] yields redundancy, contextual misinterpretation which support the previous implementations like [5] and prior methods.

Implementation of chrome-based extension requires optimization and smooth processing of information/mails, so is presenting the insights with the takeaways from [7] by reducing the API calls as fewer requests results in faster loading of report which is fundamental principle behind any system. By enabling HTTP/2 , HTTP/1.1 will reduce the latency and optimize our system in generating the reports and load faster in Chrome even on slower connections from the back-end post summarizing and prioritization.

Natural Language Processing (NLP) techniques have further enhanced email analysis by enabling text summarization and sentiment analysis. Transformer-based models such as BERT significantly improved contextual understanding in summarization tasks.

III. Ease Of Use

The Smart Mail Mentor extension is designed to enhance user experience by offering a seamless and automated approach to email management. The following features ensure ease of use:

Automated Email Classification: The system categorizes emails into predefined groups such as Promotional, Personal, Business, Order Confirmation, and Subscription Renewals without user intervention. This classification is powered by ML models like LSTM and Random Forest, ensuring high accuracy and minimal manual effort.

Simple Setup and Minimal Configuration: Users only need to authenticate the extension using the Google Gmail API. Once connected, the system automatically fetches emails and processes them without requiring further input.

Prioritization of Important Emails: The extension filters emails based on urgency, sender importance, and content using NLP techniques, ensuring that users can focus on high-priority messages first.

Concise Email Summarization: Instead of manually reading multiple emails, users receive a daily or weekly report with AI-generated summaries, significantly reducing time spent on email management.

Seamless Browser Integration: As a browser extension, Smart Mail Mentor runs in the background without disrupting user workflow. It provides real-time insights without requiring additional software installations.

By integrating AI-powered automation, Smart Mail Mentor simplifies email management while enhancing productivity, ensuring a smooth and user-friendly experience.

TABLE I
Table Type Styles

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a]Sample of a Table footnote.

Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

Acknowledgment

....

References

[1] Mansoor RAZA, Nathail Dilshani Jayasinghe and Muhana Magbouli Ali Muslam ”A Comprehensive Review On Email Spam Classification using Machine Learning Algorithms” Publisher IEEE, 2021.
[2] Palwasha Bhatti,Zunera Jalil,Asma Majeed ”Email Classification using LSTM: A Deep Learning Techniques”Publisher IEEE,2021
[3] Lorelyn F.Adrales,Ariel M Sison ”Deep Learning Approaches to a School-Email Classification With TF-IDF-based vectorization” Publisher IEEE,2024.
[4] A. S. U. R. K. S. Harish and P. S. Bhat, ”Email classification using deep learning with LSTM,” Procedia Computer Science
[5] G. U. Kiran, G. B. Desale, R. Gandi, C. U. Rao, M. Lavanya, and B. V. Reddy, “Deep Learning Based Abstractive Text Summarization: A Survey,” 2024 Parul International Conference on Engineering and Technology (PICET), pp. 1–8, 2024.
[6] K. S. Kavyashree, T. S. V, S. R, and S. R, ”Survey on Automatic Text Summarization using NLP and Deep Learning,” 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), pp. 523-527, 2023.
[7] L. Cegan, ”Improving the performance of loading web pages in the environment with high latency connections,” Dept. Inf. Technol., Fac. Elect. Eng. Informatics, Univ. Pardubice, Pardubice, Czech Republic, 2015.