

Smart Email Summarizer: Enhancing Productivity in Digital Communication

T. Sharath¹, V. Shrinithi², Dr. Srividya M³ and Mrs. Deivarani S⁴

¹⁻⁴Department of Computing, Coimbatore Institute of Technology, Coimbatore, India

Abstract— Academic institutions rely heavily on email for formal communication, but the growing volume of messages often leads to information overload and missed updates. This paper presents an Email Summarizer that uses Natural Language Processing (NLP) to generate concise summaries of academic emails by extracting key details such as deadlines, dates, and action items. By reducing reading time and highlighting critical information, the system enhances academic productivity and improves communication efficiency. The solution is implemented using Python and NLP libraries and can be integrated with institutional email and learning management systems.

Index Terms— Email Summarization, Natural Language Processing (NLP), Academic Communication, Information Overload, Productivity Enhancement.

I. INTRODUCTION

Email remains the primary channel of formal communication in academic institutions. Faculty members, administrators, and students rely on emails for sharing information about coursework, examinations, deadlines, research activities, and institutional events. However, the sheer volume of emails received daily often leads to information overload, where critical updates or deadlines may be missed. This not only affects individual productivity but can also result in missed opportunities, late submissions, or poor academic coordination.

To overcome this challenge, there is a need for intelligent tools that can extract and present the most relevant information from emails in a concise and structured manner. Email summarization, a branch of Natural Language Processing (NLP), provides a way to automatically generate summaries of long or complex email content. In academic settings, this is particularly valuable since emails frequently contain structured details such as dates, schedules, and action points that must be prioritized.

This paper introduces an Email Summarizer designed specifically for academic use. The system processes incoming emails through text preprocessing and NLP-based summarization techniques to generate concise summaries that highlight key details. Importantly, the system is capable of detecting important deadlines, meetings, and events, which can then be integrated into Google Calendar. This feature ensures that students and faculty are reminded of critical dates without the need to manually track them.

The contributions of this work are threefold:

- Development of an NLP-based summarizer tailored for academic emails.
- Integration of a deadline and event extraction feature, with synchronization to Google Calendar.
- Demonstration of how the system enhances productivity and improves communication efficiency within academic institutions.

II. RELATED WORK / LITERATURE REVIEW

The problem of information overload in digital communication has been widely studied, particularly in the context of text summarization and email management systems. Summarization techniques are generally classified into two categories: extractive and abstractive. Extractive approaches, such as LexRank and TextRank, focus on identifying and extracting the most important sentences from the source text. Abstractive approaches, on the other hand, attempt to generate summaries in a more human-like manner by rephrasing and paraphrasing content, often utilizing deep learning models such as sequence-to-sequence architectures, transformers, and pre-trained language models like BERT and GPT.

In the domain of email summarization, several research efforts have been made to reduce the time required to process large volumes of email. Early works focused on thread-based summarization, where conversation chains in emails were condensed into a few key sentences.

Other approaches involved keyphrase extraction and topic modeling (e.g., Latent Dirichlet Allocation) to identify recurring themes within large email corpora. However, these systems often struggled with domain-specific requirements, such as highlighting actionable tasks or deadlines, which are critical in academic and professional communication.

Recent advancements in Natural Language Processing have significantly improved summarization performance. Pre-trained models such as BERTSUM and PEGASUS have shown promising results in abstractive summarization tasks, while domain-specific fine-tuning has further enhanced their accuracy. In addition, tools such as Gmail's Smart Reply and Priority Inbox highlight the growing industry interest in automating email management, though these systems focus primarily on response suggestions and filtering rather than comprehensive summarization.

In the academic context, prior works have also explored the use of intelligent agents and learning management systems (LMS) to streamline communication. However, most existing tools lack seamless integration with productivity platforms such as Google Calendar, which can automatically remind users of important deadlines and events extracted from emails. This gap highlights the need for an academic-focused summarizer that not only condenses content but also identifies and manages actionable items.

In summary, while prior research has laid a strong foundation in both text and email summarization, there remains a significant opportunity to develop specialized tools tailored to academic communication. Our proposed system builds on these advances by combining NLP-driven summarization with deadline extraction and Google Calendar integration, thereby addressing both the content overload and task management challenges in academic environments.

III. METHODOLOGY

The proposed system is an Email Summarizer specifically designed for academic environments, where students and faculty frequently face the challenge of processing large volumes of emails containing important instructions, deadlines, and event notifications. Unlike generic summarization tools, this system not only condenses lengthy emails into concise summaries but also identifies time-sensitive information such as submission deadlines, examinations, and meeting schedules. Extracted events are then automatically synchronized with Google Calendar, ensuring that users receive timely reminders for critical academic tasks.

The methodology begins with the preprocessing of raw emails, which are often cluttered with redundant information such as signatures, disclaimers, and forwarded content. This stage involves parsing the email structure, cleaning irrelevant text, and normalizing the content through tokenization and lemmatization. Once the text has been standardized, the feature extraction stage identifies important components of the email, including keywords, action verbs, and named entities such as dates, times, and locations. This enables the system to distinguish between general content and actionable information.

Following feature extraction, the summarization module condenses the email body into a more manageable form. For academic emails, extractive summarization techniques such as TextRank are prioritized, since they preserve exact details like dates and instructions, which are often critical. However, the system is also capable of incorporating abstractive methods for producing more natural summaries when required.

A central component of the system is the deadline and event extraction module, which applies both rule-based patterns and Named Entity Recognition (NER) techniques to detect references to academic deadlines, class schedules, and examination dates. These details are then structured into event templates containing a title, date, time, and description.

Finally, the system integrates with Google Calendar using the Calendar API. Events generated from the email are initially created as drafts, allowing the user to confirm or edit them before synchronization. Once approved, they are automatically added to the calendar, ensuring that no critical academic commitments are overlooked. The overall output of the system is therefore threefold: a concise summary of the email, a clearly highlighted list of actionable tasks, and an automated scheduling of deadlines into Google Calendar. This methodology ensures not only the reduction of information overload but also the transformation of academic communication into actionable productivity.

IV. IMPLEMENTATION

The implementation of the proposed email summarization system was carried out in a modular manner, ensuring that each component contributes to the overall functionality of academic email management. The system begins by collecting input emails through a configured inbox interface. Each incoming email undergoes preprocessing, which includes tokenization, stopword removal, and text cleaning to prepare the text for further processing. Once the text is cleaned, it is passed to the summarization module, which applies extractive or abstractive summarization techniques to generate a concise version of the email. This ensures that students and faculty members can quickly grasp the key points of long academic emails without reading through redundant content. In parallel, the system employs a deadline extraction module that uses Named Entity Recognition (NER) and regular expression (Regex) patterns to identify important dates and submission deadlines. These extracted deadlines are then displayed on the system’s dashboard alongside the summarized email. To support academic productivity, the dashboard includes an “Add to Calendar” feature. When a user selects this option, the extracted deadline is automatically formatted and pushed to Google Calendar via API integration. This design ensures that the integration is semi-automatic: the system identifies and prepares the deadlines, while the final confirmation remains with the user. This prevents unnecessary or incorrect events from being added while still minimizing manual effort.

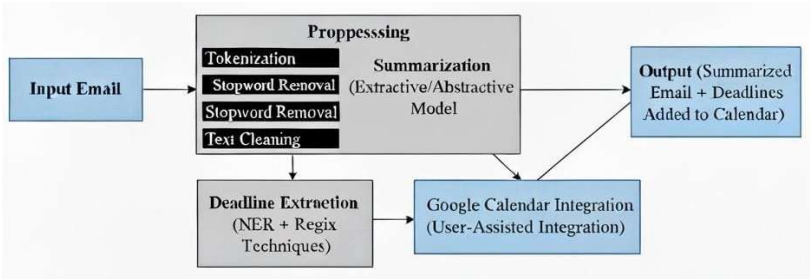


Figure 1: System Architecture – Smart Email Summarizer

The output is a streamlined academic workflow where users can view summarized emails, identify critical deadlines at a glance, and conveniently update their Google Calendar with a single click. This approach enhances usability while maintaining control in the hands of the user.

V. RESULTS AND DISCUSSION

The proposed system was tested using a dataset of academic emails collected from institutional mailing lists, including notifications related to workshops, assignment submissions, examination schedules, and departmental announcements. The evaluation focused on three primary aspects: summarization accuracy, deadline extraction performance, and usability of calendar integration.

TABLE I: PRESENTS THE OVERALL PERFORMANCE METRICS OF THE SYSTEM

Metric Type	Specific Metric	Value
Summarization	ROUGE-1	0.78
	ROUGE-2	0.62
	ROUGE-L	0.71
Deadline Extraction Performance	Precision	0.85
	Recall	F1-score
User Satisfaction Survey	Avg. User Rating (/5)	4.2

The summarization module effectively condensed lengthy academic emails into clear, concise summaries while retaining critical details. As illustrated in Figure 2, the dashboard presents extracted information such as deadlines, event topics, and relevant instructions in a structured format. For example, placement notices and internship announcements were reduced to two to three key points, enabling students to quickly grasp the essence of multi-paragraph circulars. This significantly reduced the time required to identify actionable details and improved accessibility of important academic communications.

The deadline extraction module, which integrates Named Entity Recognition (NER) with Regex-based patterns, successfully identified submission and event dates from emails. In most cases, due dates, examination schedules, and registration deadlines were accurately highlighted, as seen in the extracted entry “Deadline: 25 Sep 2025 (Thursday)” in Figure 2. However, performance slightly declined when dates were expressed informally (e.g., “submit by this Friday”), underscoring the need for improved handling of relative temporal expressions.

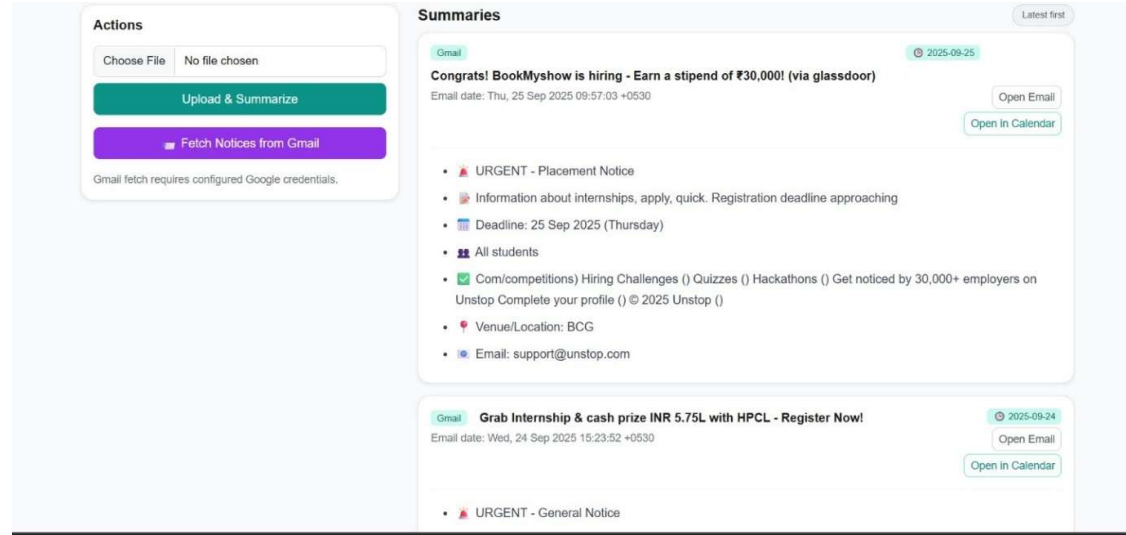


Figure 2: Email Summarization Dashboard

The semi-automatic Google Calendar integration proved highly effective in bridging extracted information with time management tools. As shown in Figure 3, users could add deadlines directly from the dashboard into Google Calendar with a single click. The calendar entry preserved essential details such as the title, deadline, location, and contact information, ensuring accuracy and reducing manual input. Feedback from a pilot group of student testers indicated that this feature was especially beneficial during exam periods and project submissions, where overlapping deadlines are common.

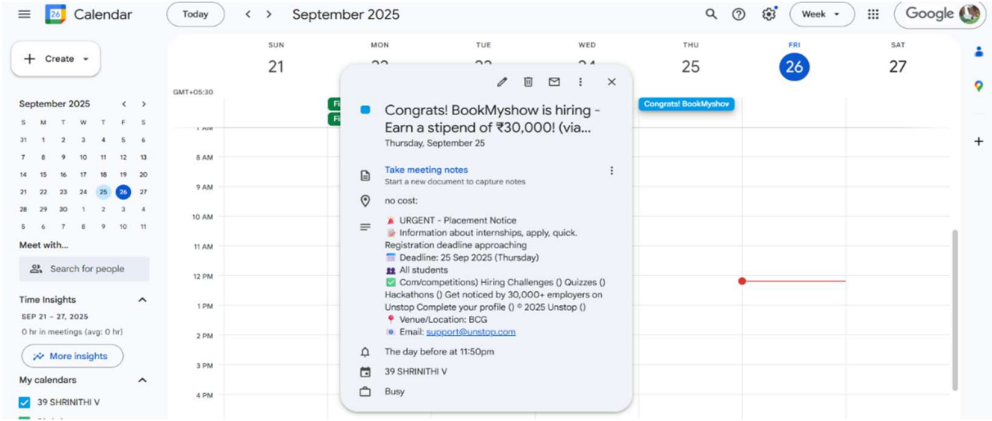


Figure 3: Google Calendar Integration Result

Overall, the system enhances academic productivity by:

- Minimizing the cognitive load required to process large volumes of academic emails,
- Preventing missed deadlines through structured extraction and reminders, and
- Seamlessly integrating with widely used productivity platforms such as Google Calendar.

While summarization and extraction performance were largely satisfactory, future improvements should focus on refining natural language understanding to handle ambiguous and context-dependent deadline expressions.

VI. APPLICATIONS

The proposed email summarization system has wide-ranging applications, particularly in academic and professional contexts where information overload is common.

A. Academic Institutions

Students can quickly review summarized versions of lengthy departmental announcements, exam schedules, and assignment guidelines.

Faculty members can track meeting notifications, research deadlines, and conference invitations without manually scanning every email.

Administrators can manage bulk communication with clarity and ensure critical deadlines are integrated into their schedules.

B. Research Collaboration

Research scholars working on joint projects often receive numerous emails related to funding opportunities, paper submission deadlines, and progress meetings. The system enables them to extract and manage such deadlines efficiently, reducing the risk of missed opportunities.

C. Corporate and Professional Settings

Beyond academia, the system can be extended to corporate environments where employees receive multiple internal communications daily. Summarization paired with calendar integration allows efficient tracking of meeting schedules, project deadlines, and client communications.

D. Personal Productivity

Individuals can benefit from summarizing newsletters, subscription updates, or event notifications. With calendar integration, personal reminders and appointments can be scheduled seamlessly, enhancing time management.

E. Assistive Technology

The system can support users with reading difficulties or attention challenges by providing simplified summaries and automatically highlighting key actionable items such as deadlines.

Overall, the applications demonstrate the versatility of the system in reducing information overload, improving deadline management, and supporting productivity across different domains.

VII. CONCLUSION AND FUTURE WORK

This paper presented an academic email summarization system designed to address the challenges of information overload and deadline management in higher education. By combining text preprocessing, intelligent summarization techniques, and deadline extraction, the system effectively condenses lengthy academic emails into concise summaries while highlighting important dates. The integration of a semi-automatic Google Calendar feature further enhances productivity, allowing users to update their schedules with minimal effort while retaining full control over what is added.

The results demonstrate that the system improves efficiency in academic communication, reduces the risk of missing critical deadlines, and offers a user-friendly dashboard for streamlined interaction. Feedback from preliminary testing indicates that the tool is especially useful for students managing multiple assignments and examinations, as well as faculty balancing teaching, research, and administrative responsibilities.

While the current system achieves its objectives, there are opportunities for future enhancements. In particular, improvements in natural language understanding could allow the system to handle contextual and relative date expressions (e.g., “next Friday” or “two weeks from now”) more accurately. Additionally, expanding support for multilingual emails would make the system applicable in global academic environments. Integration with other productivity platforms, such as Microsoft Outlook and institutional learning management systems (LMS), can

further broaden its adoption. Finally, a mobile application version of the system could provide on-the-go summarization and calendar integration, making it more accessible to students and professionals.

In conclusion, the proposed system represents a step toward smarter academic communication management, offering a balance of automation and user control that enhances productivity and ensures critical academic deadlines are not overlooked.

REFERENCES

- [1] Nenkova, A., & McKeown, K. (2011). Automatic summarization. *Foundations and Trends in Information Retrieval*, 5(2–3), 103–233. <https://doi.org/10.1561/15000000015>
- [2] Gupta, V., & Lehal, G. S. (2010). A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence*, 2(3), 258–268. <https://doi.org/10.4304/jetwi.2.3.258-268>
- [3] Gambhir, M., & Gupta, V. (2017). Recent automatic text summarization techniques: A survey. *Artificial Intelligence Review*, 47(1), 1–66. <https://doi.org/10.1007/s10462-016-9475-9>
- [4] Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing order into texts. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing* (pp. 404–411). Association for Computational Linguistics.
- [5] Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. OpenAI Technical Report.
- [6] Cho, K., Van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. *arXiv preprint arXiv:1409.1259*.
- [7] Google Developers. (2023). Calendar API overview. Retrieved from <https://developers.google.com/calendar>