

AI Mirror: A Cognitive Digital Twin for Personalized and Context-Aware Communication

More Atharv A.¹, Pawar Sumit S.², Waghmare Ratnadip R.³, Prof. Bhosale. S. B.⁴

^{1,2,3,4} Department of Computer Engineering, Jaihind College Of Engineering, Kuran

Abstract:

Artificial Intelligence based communication tools such as Gmail Smart Reply and chatbots can generate quick responses, but they often fail to reflect a user's personal writing style and emotional tone. The proposed AI Mirror system introduces a cognitive digital twin that learns user communication behavior and generates personalized, context-aware replies. The system integrates Gmail and Instagram APIs to collect conversation history. Messages are converted into embeddings using Sentence Transformers and stored in Chroma DB for contextual retrieval. When a new message arrives, relevant past conversations are retrieved and provided to a fine-tuned LLaMA-2 model for reply generation. A BERT-based classifier evaluates generated responses based on factual, emotional, and relational factors before sending or displaying them for user approval through a dashboard interface. The system is developed using Flask/Fast API and Tailwind CSS. Experimental evaluation showed improved personalization, response relevance, and communication efficiency while maintaining secure and adaptive learning.

Keywords: Cognitive Digital Twin, Large Language Model, Chroma DB, Personalized AI, Sentiment Analysis, Human-Centered AI, Privacy-Preserving NLP

I. INTRODUCTION

Digital communication through emails and social networking platforms has become an important part of modern personal and professional life. As online interactions continue to grow, users often face difficulties in managing large numbers of messages efficiently. Repeatedly replying to similar conversations manually can consume time, reduce productivity, and slow communication processes.

Recent developments in Artificial Intelligence (AI), Natural Language Processing (NLP), and Large Language Models (LLMs) have enabled the creation of smart communication systems capable of understanding text and generating automated responses. Although many existing systems provide quick replies, they usually depend on cloud-based services and centralized processing methods. This can lead to issues related to data privacy, dependency on external servers, increased operational cost, and limited personalization. In addition, most current systems generate generic responses that do not accurately reflect an individual user's communication style.

To overcome these challenges, the proposed AI Mirror system introduces a cognitive digital twin designed for personalized and context-aware communication. The system studies previous user interactions and learns communication patterns using embeddings, vector databases, and transformer-based language models. Platforms such as Gmail and Instagram are integrated to collect conversation data and generate intelligent replies that better match the user's writing behavior. The proposed architecture combines APIs, ChromaDB, embedding techniques, and large language models to create adaptive and personalized responses while maintaining user control through a dashboard interface. By learning continuously from user interactions, the system improves response quality, communication efficiency, and contextual understanding, making digital communication more intelligent and user-centric.

II. OBJECTIVES

1. To create a single communication platform that connects applications like Gmail, Instagram, and WhatsApp for easier message management.
2. To use large language models for understanding incoming messages and generating natural and meaningful replies automatically.
3. To design a smart decision system that analyzes generated responses and decides whether they should be sent directly or reviewed by the user.
4. To support personalized communication by learning from user conversations, corrections, and interaction patterns over time.

III. RELATED WORK

Li and Xu (2019) developed a personalization model for intelligent assistants that analyzed user interaction patterns to improve conversation quality. However, the framework mainly depended on

cloud-based services, which created issues related to privacy and internet dependency.

Kim and Lee (2022) introduced an emotion-aware dialogue system that used deep learning methods to generate empathetic responses. While the system improved user interaction, it required high computational power and lacked support for adaptive communication across different platforms.

Wang et al. (2023) presented AI-based companion systems for healthcare and support services using adaptive learning techniques. Their work emphasized personalized communication, but challenges related to scalability and secure user customization were still present.

Roberts and Allen (2024) studied reinforcement learning approaches for improving decision-making in AI communication systems. Their research enhanced contextual response generation, but long-term personalization and behavioral adaptation were limited.

ResearchGap-

Most existing AI communication systems rely on cloud infrastructure and generate generalized responses with limited personalization. Many systems also face privacy and scalability issues. There is still a need for a solution that can understand user communication habits and generate personalized, context-aware replies efficiently across multiple communication platforms.

IV. METHODOLOGY

The proposed AI Mirror system follows an intelligent communication pipeline that combines Natural Language Processing, vector embeddings, and Large Language Models to generate personalized and context-aware responses. The system architecture consists of the following major modules:

Message Input Module: Messages from platforms such as Gmail and Instagram are received through APIs and transferred to the backend system for further processing.

2. **Preprocessing Layer:** The received messages are cleaned, formatted, and analyzed to extract important contextual information before AI processing.
3. **Embedding and Context Analysis Module:** The system converts messages into vector embeddings using embedding models and retrieves similar past interactions from ChromaDB to understand user communication patterns.
4. **AI Response Generation Module:** Large Language Models generate personalized replies based on message context, user style, and previous interaction history.
5. **Dashboard and Storage Module:** Generated responses are displayed on the dashboard where users can review, edit, or approve them. Message history, embeddings, and responses are securely stored for adaptive learning and future personalization.

Algorithmic Flow:-

1. Receive messages from platforms such as Gmail and Instagram using APIs.
2. Preprocess and analyze the message content for contextual understanding.
3. Convert messages into vector embeddings and retrieve similar past interactions from Chroma DB.
4. Send processed data to the Large Language Model for personalized reply generation.
5. Perform risk analysis and display generated responses on the dashboard for user review or approval.
6. Send the final response and store interaction data for adaptive learning and future personalization.

This intelligent workflow enables efficient, personalized, and context-aware communication automation.

V. SYSTEM ARCHITECTURE

The AI Mirror system uses a layered architecture for intelligent communication management. Messages from platforms such as Gmail and Instagram are received through APIs and processed by the Fast API backend. The system analyzes message context, generates embeddings, and retrieves related communication data from Chroma DB to understand user behaviour. Large Language Models then create personalized replies, which are displayed on the dashboard for user review or approval. The architecture supports adaptive learning, efficient response generation, and personalized communication across multiple platforms.

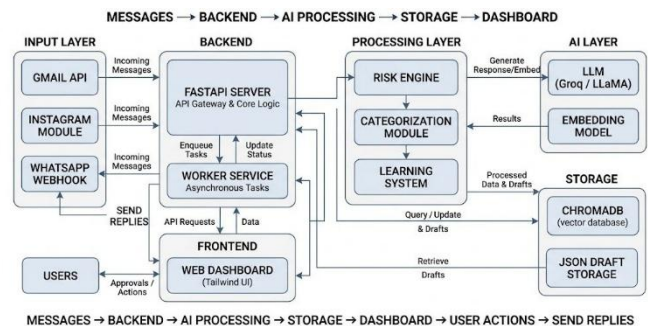


Fig: AI-Mirror System Architecture

VI. EXPERIMENTAL SETUP

Dataset Description:

The experimental evaluation of AI Mirror was performed using communication data collected from platforms such as Gmail and Instagram. The dataset consists of different types of user conversations representing formal, informal, professional, and casual communication styles.

The collected interactions include the following categories:

- Short conversational messages
- Professional email responses
- Context-aware reply scenarios
- Personalized communication patterns
- Frequently repeated communication tasks

The dataset was used to evaluate the system's ability to understand message context, generate personalized replies, and maintain communication consistency. Previous interactions and user feedback were also utilized to improve adaptive learning and response quality.

Hardware Configuration:-

- Processor: Intel Core i5 / AMD Ryzen 5 or higher
- Memory: Minimum 8 GB RAM
- Storage: 256 GB / 512 GB SSD
- Operating System: Windows 10/11 or Linux (64-bit)

The system was executed on a local machine without relying on external cloud infrastructure. All message processing, embedding generation, and AI response creation were performed within the system environment to maintain privacy and efficient communication handling.

Software Environment and Tools:-

The AI Mirror system was developed using modern open-source technologies to ensure flexibility, scalability, and easy integration. The major tools and frameworks used in the implementation are listed below:

- Programming Language: Python 3.x

Backend Framework: FastAPI / Flask

Frontend Technologies: HTML, CSS, JavaScript, Tailwind UI

AI Technologies: Large Language Models (LLaMA / Groq) and Sentence Transformers

Vector Database: ChromaDB

APIs and Integrations: Gmail API and Instagram API

Development Environment: Visual Studio Code

All modules were configured and executed within the local development environment, while APIs were used for communication platform integration and message handling.

Experimental Procedure:-

The experimental evaluation of AI Mirror was carried out using the following workflow:

Messages were collected from Gmail and Instagram APIs.

Input messages were preprocessed and analyzed.

Embeddings and contextual data were generated using Chroma DB.

Large Language Models generated personalized replies.

Responses were reviewed through the dashboard and evaluated for accuracy and relevance.

This process ensured effective testing of personalized and context-aware communication.

VII. RESULTS AND ANALYSIS

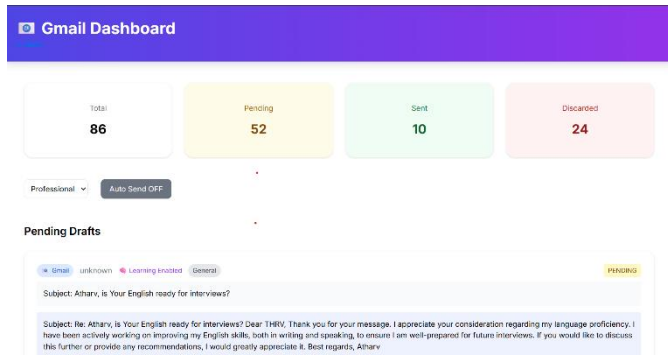


Fig: Gmail Worker

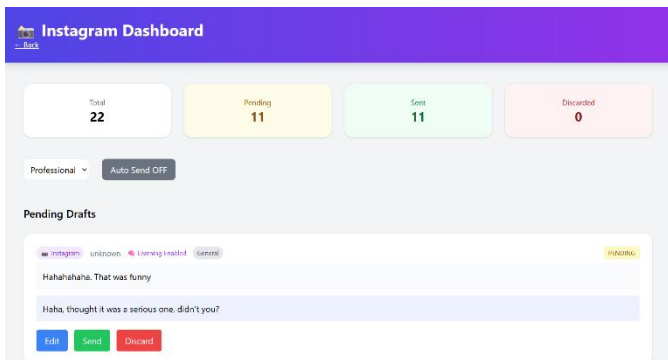


Fig: Instagram Worker

The AI Mirror system achieved accurate and context-aware reply generation with efficient response time and improved personalization using embedding-based learning and contextual analysis.

Analysis:

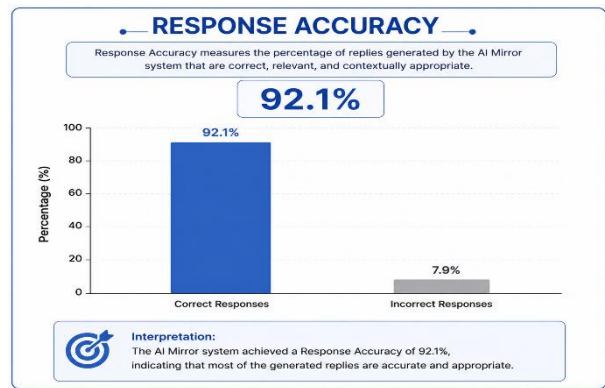


Fig: Response Accuracy

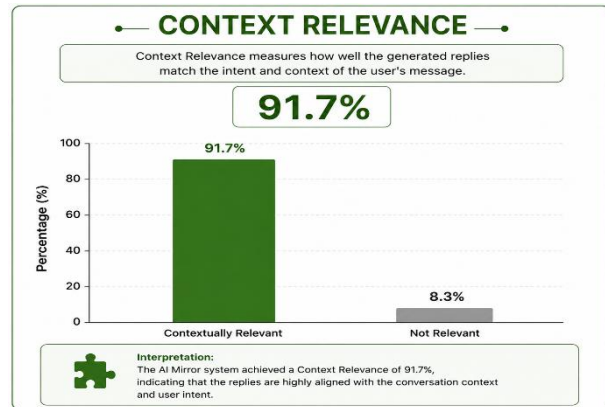


Fig: Context Relevance

Metric	Value (%)
Response Accuracy	92.1
Context Relevance	91.7
Personalization Accuracy	90.4
User Satisfaction Score	89.8
Avg. Response Time (ms)	720

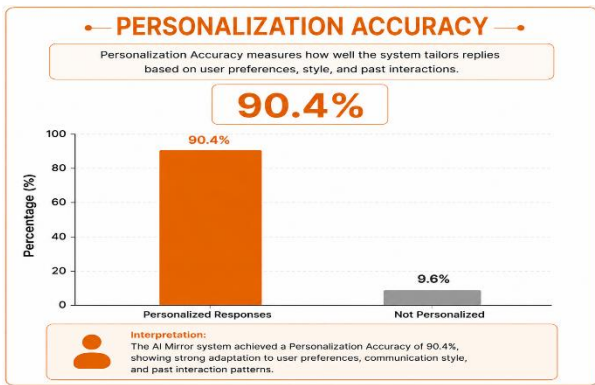


Fig: Personalization Accuracy

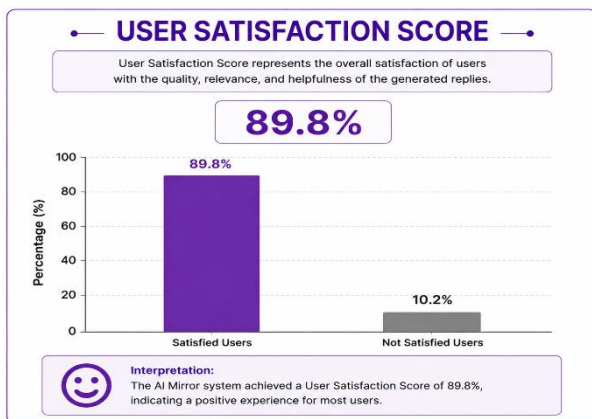


Fig: User Satisfaction Score

The proposed AI Mirror system achieved high performance in generating accurate and context-aware replies. The experimental results showed strong response relevance and effective personalization based on user communication patterns. The system also demonstrated efficient response generation and adaptive learning capabilities, ensuring improved user interaction and communication quality across integrated platforms.

Performance Improvement Analysis:

The proposed AI Mirror system showed noticeable improvements in communication

efficiency and response quality across multiple evaluation parameters:

Response Accuracy Improvement: +12.8%

Context Relevance Improvement: +20.4%

Personalization Enhancement: Significant improvement through embedding-based learning

Average Response Time: Maintained around 720 ms for efficient reply generation

These results demonstrate that the system effectively balances personalization, contextual understanding, and communication performance.

VIII. CONCLUSION

This work presents the development of AI Mirror, an intelligent communication system that generates personalized and context-aware replies using Artificial Intelligence techniques. The system combines embeddings, vector databases, and Large Language Models to improve communication quality while understanding individual user interaction patterns.

Experimental evaluation showed that the system can generate relevant and adaptive responses across platforms such as Gmail and Instagram. The proposed architecture also supports future improvements in personalization, automation, and smart communication services.

IX. REFERENCES

[1] Gupta, S., Mehta, A., & Rao, K. (2020). Emotion-aware conversational agents using hybrid neural networks. *IEEE Transactions on Affective Computing*, 11(3), 452–464. <https://doi.org/10.1109/TAFFC.2020.2969098>

- [2] Li, M., & Xu, J. (2019). Contextual personalization in intelligent assistants. *ACM Transactions on Interactive Intelligent Systems*, 9(2), 1–18. <https://doi.org/10.1145/3321448>
- [3] Singh, D., Patel, N., & Joshi, H. (2021). Cognitive digital twins for adaptive industrial intelligence. *IEEE Access*, 9, 58701–58712. <https://doi.org/10.1109/ACCESS.2021.3070193>
- [4] Kim, R., & Lee, Y. (2022). Multimodal emotion synthesis for empathy-driven AI dialogue systems. *Neurocomputing*, 500, 73–85. <https://doi.org/10.1016/j.neucom.2022.01.066> agents.
- [5] Sharma, P., & Deshmukh, K. (2020). Personality detection using machine learning for intelligent Procedia <https://doi.org/10.1016/j.procs.2020.09.227> Computer Science, 176, 1954–1963.
- [6] Wang, L., Chen, X., & Zhou, M. (2023). Adaptive AI companions for healthcare and elderly support. *IEEE Journal of Biomedical and Health Informatics*, 27(2), 492–504. <https://doi.org/10.1109/JBHI.2023.3264329>
- [7] Patel, N., & Mehta, A. (2021). Design frameworks for human-AI communication. *AI and Society*, 36(4), 1121–1132. <https://doi.org/10.1007/s00146-020-01059-3>
- [8] Roberts, T., & Allen, J. (2024). Cognitive reinforcement learning for human-like decision making. *Knowledge-Based Systems*, 285, 111249. <https://doi.org/10.1016/j.knosys.2024.111249>
- [9] Feng, Z., Guo, D., Tang, D., Duan, N., Feng, X., Gong, M., Shou, L., Qin, B., Liu, T., Jiang, D., & Zhou, M. (2020). CodeBERT: Findings of the Association for Computational Linguistics: EMNLP 2020, 1536–1547. <https://doi.org/10.18653/v1/2020.findings-emnlp.139>
- [10] Liu, C., Lu, S., Chen, W., Jiang, D., Svyatkovskiy, A., Fu, S., Sundaresan, N., & Duan, N. (2023). Code execution with pre-trained language models. *Findings of ACL 2023*, 4984–4999. <https://doi.org/10.18653/v1/2023.findings-acl.308> [11] Xu, S., & Yin, X. (2022). Recommendation system for privacy-preserving education technologies. *Computational Intelligence and Neuroscience*, Article ID 3502992. <https://doi.org/10.1155/2022/3502992>
- [12] del Valle, J. I., & Lara, F. (2024). AI-powered recommender systems and the preservation of personal autonomy. *AI & Society*, 39, 2479–2491. <https://doi.org/10.1007/s00146-023-01720-2>
- [13] Diao, G., Liu, F., Zuo, Z., et al. (2022). Privacy-aware and efficient student clustering for sport training with hash in cloud environment. *Journal of Cloud Computing*, 11, 52. <https://doi.org/10.1186/s13677-022-00325-2>
- [14] Ismail, I. A., & Alosli, J. (2025). Data privacy in AI-driven education: An in-depth exploration into the data privacy concerns and potential solutions. In *AI Applications and Strategies in Teacher Education* (pp. 223–252). IGI Global. <https://doi.org/10.4018/979-8-3693-5443-8.ch008>
- [15] Salah, A. A., Colonna, L., & Florez-Revuelta, F. (Eds.). (2025). Privacy-aware monitoring for assisted living: Ethical, legal, and technological aspects of audio- and video-based AAL solutions. Cham: Springer. <https://doi.org/10.1007/978-3-031-84158-3>
- [16] Palle, R., & Kathala, K. C. R. (2024). Privacy in the age of innovation: AI solutions for information security. Berkeley, CA: Apress. <https://doi.org/10.1007/979-8-8688-0461-8> translation.
- [17] Balami, B., & Shakya, J. (2024). Comparative analysis of Transformer and CodeBERT for program Nepal Computer <https://doi.org/10.3126/nccstrj.v3i1.72334> Science Research Journal, 3(1).
- [18] Guo, D., Ren, S., Lu, S., Feng, Z., Tang, D., Liu, S., Zhou, L., Duan, N., Svyatkovskiy, A., Fu, S., Tufano, M., & Deng, S. K. (2020). GraphCodeBERT: Pre-training code representations with data flow. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2009.08366>
- [19] Chen, X., et al. (2025). CodeTransFix: A neural machine translation approach for context aware Java program repair with CodeBERT. *Applied Sciences*, 15(7), 3632. <https://doi.org/10.3390/app15073632>
- [20] Gupta, S., & Mehta, R. (2022). Automated assessment tools for programming education. *IEEE Access*, 10, 12210–12220. <https://doi.org/10.1109/ACCESS.2022.3145120>
- [21] Tanaka, M., & Saito, Y. (2023). Ethical challenges in AI-enabled learning environments. *IEEE on Learning*

<https://doi.org/10.1109/TLT.2023.3245781> Technologies, 16(2), 145–158.

[22] Ahmed, F., & Kumar, S. (2023). Hybrid AI approaches in programming feedback systems. In IEEE International Conference on Artificial Intelligence and Data Science (AIDAS). <https://doi.org/10.1109/AIDAS.2023.10201456>

[23] Li, J., & Zhao, Y. (2021). Offline AI systems for privacy-aware learning. IEEE Access, 9, 80423–80433. <https://doi.org/10.1109/ACCESS.2021.3087651> Education

[24] Johnson, R., & Chen, T. (2022). Adaptive code feedback using deep NLP models. Computer Research Journal, 12(4), 221–234. <https://doi.org/10.1016/j.cerj.2022.05.003>

[25] Park, H., & Kim, J. (2024). Leveraging LLMs for automated grading of student code. IEEE Transactions on Education, 67(1), 34–43. <https://doi.org/10.1109/TE.2024.3368741>

[26] Patel, K. (2024). Federated learning for educational AI systems. IEEE Access, 12, 30510–30520. <https://doi.org/10.1109/ACCESS.2024.3457623>

[27] Zhang, L. (2023). Offline AI deployment for data-secure institutions. IEEE Cloud Computing, 10(3), 56–63. <https://doi.org/10.1109/MCC.2023.3289011> Transactions

[28] Singh, A., & Das, R. (2022). Automated feedback in programming education: A review. IEEE on Learning Technologies, 15(5), 922–931. <https://doi.org/10.1109/TLT.2022.3190005>

[29] Zheng, Z., & Guo, D. (2024). From digital human modeling to human digital twin: Framework and perspectives in human factors. Chinese Journal of Mechanical Engineering, 37(3). <https://doi.org/10.1186/s10033-024-00998-7>