National College of Ireland

The Role of AI Chatbots in Enhancing Efficiency and Customer Experience in Banking Call Centers

MSc Research Project
Programme Name: MSCFTD

Forename Surname: Qingyun Cai
Student ID: 23403519

School of Computing
National College of Ireland

Supervisor: Victor del Rosal

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Qingyun Cai		
Student ID:	23403519		
Programme:	MSCFTD	**Year:**	2024/2025
Module:	Practicum		
Supervisor:	Victor del Rosal		
Submission Due Date:	11/08/2025		
Project Title:	The Role of AI Chatbots in Enhancing Efficiency and Customer Experience in Banking Call Centres		
Word Count:	6328	**Page Count:**	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Qingyun Cai

Date: 11/08/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

The Role of AI Chatbots in Enhancing Efficiency and Customer Experience in Banking Call Centers

Forename Surname Qingyun Cai
Student ID 23403519

Abstract

The infusion of AI into the Banking sector has had a substantial influence on customer service, notably with the emergence of AI Chatbots. This research initiative focused on the creation, deployment, and evaluation of PrimeBank AI Assistant, a Rule-Based Banking Chatbot developed to enhance efficiency and optimize customer experience in Digital Banking Services. The chatbot has been developed using standard web technologies like HTML, CSS and JavaScript on a modular and Client-side architecture ensuring that the chatbot is secure, responsive and does not store sensitive information about the users.

The aim of overarching this project was to learn about and apply Rule Based Chatbot Systems that can mimic the real-world application and can support Banking and support users with relevant questions as well as usability on any device. We compiled a 1000 Q&A Pair Dataset to train and better enable our chatbot's logic. We used three real-world case studies to test its capabilities against 3 important and common challenges in chatbot design: conversation context, culturally/emotionally sensible responses and Regulatory Compliance with multi-chat session history. Iterative improvements to the design were made to effectively solve each challenge and demonstrate the chatbot's growing ability.

The result of the project shows that the basic rule-based AI assistants can be improved using the concepts of structured logic and user-oriented design approaches. In addition, the success of the chatbot in this project also opens the doors for more seamless integration with NLP (Natural Language Processing) framework in the future, to have more automated and smarter support and financial solutions. In this report, we have discussed methodology, design, implementation and case study-based evaluations to explore various ways to build secure, scalable and context aware chatbots for the banking industry.

1 Introduction

AI is being increasingly used to transform how customer service is being delivered in banking. Chatbots are now being considered as an alternative and scalable support channel over traditional ones. An increased demand for customer interactions to be instant, precise and accessible has also made banks start looking for alternative AI-based support that abides by privacy and regulatory demands.

The goal of this project is to create and evaluate the performance of a rule-based chatbot called PrimeBank AI Assistant, a chatbot capable of simulating bank interactions by responding to frequently asked questions. The chatbot is developed using standard web technologies (HTML, CSS, JavaScript) and is deployed entirely by client-side to ensure user privacy. The chatbot is trained on a dataset of 1,000 question-answer pairs to provide fast, context-aware, and reliable responses to typical customer queries. The system also includes multiple web tools (calculator, branch locator, etc.) to augment the conversation.

The project is closely related to the real-world banking processes. A set of case studies has been designed to evaluate the system under multiple challenging scenarios such as sustaining the context of multi-turn conversations, checking emotional, Regulatory Compliance and cultural sensitivity of the generated responses, handling multiple chat sessions in parallel. Case studies also help to understand the limitations of rule-based logic and the scope of future enhancements.

This paper is organized in the following sections. In the first section is a literature review to find related research and technologies. Methodology section is describing data gathering and logic design process. Design and Implementation parts describe chatbot's structure and development workflow. After that, Evaluation and Discussion part provides practical testing and future work. Finally, this project will be able to prove that even with the rudimentary AI logic, well-designed chatbot has the potential to make banking support service of better quality and availability.

2 Related Work

This chapter reviews contemporary studies on AI chatbot applications in bank call centers. It is organized along three themes: factors contributing to performance and end-user adoption (Section 2.1); chatbot design requirements related to emotional intelligence, cultural and regional sensitivities, and legal concerns (Sections 2.2–2.5); and an integrative synthesis of critical research gaps informing the design and implementation of the PrimeBank AI Assistant (Section 2.6).

2.1 Chatbots in Banking: Use Cases and Industry Evolution

AI chatbots are a key feature of current banking systems, as they have a range of customer-facing applications and help to support banking services, such as those available in call centers, with scalable, low-cost and always-available automation for standard questions. Cîmpeanu et al. (2023) provide a comprehensive overview of chatbot uses in banks, including handling transactions, instant query answering and other informational services like balance checks and product information. In particular, these low-value, high-frequency interactions are the target for the PrimeBank AI Assistant, and there are some clear overlaps with the specific use-cases the research team is focusing on.

The decision to use chatbots is not based on efficiency considerations alone. User behavior and experience are also fundamental. Saxena et al. (2023) and Srivastava et al. (2024) point out that customers' perception of trustworthiness, ease of use, and usefulness are the best predictors of their adoption. These papers suggest that a good user interface (UI) should be transparent and simple as well, and that has been part of the UI development process for PrimeBank from the start.

Widjaja and Legowo (2024) used an Indonesian bank's customers to study the context accuracy and the impact of latency on their overall satisfaction. It was found that clients are much more satisfied when the chatbot's replies accurately reflect the queries' intent and when those replies are given with a minimal delay. These results explain the choice of a rule-based logic engine and a curated dataset for PrimeBank, which allow for both low latency and a high degree of semantic accuracy.

Customer retention is another theme in the selected literature. For instance, Park et al. (2024) demonstrated that return usage is directly related to "structured navigation, clear information presentation, and quick response times" (p. 196). PrimeBank's design is guided by these results, such as using category-based menus and concise, tiered responses, to minimize frustration and navigation fatigue.

Salem (2024) investigated how chatbots can expand banking services to underbanked and rural populations. The results indicate that modular chatbot design allows for quick adaption to local context, including language and connectivity limitations. PrimeBank Assistant was designed to embody this adaptability by being functional in a browser offline mode and having a modular content structure for easy updates.

The issue of security is also still important. Chanda and Prabhu (2023) and Kovacevic et al. (2024) state that without privacy by design, AI systems could create new security risks, such as the potential for data leakage or unauthorized access. PrimeBank follows a privacy-by-design approach, operating only on the client's side with no data storage or remote processing.

In conclusion, Mahmud and Uddin (2023) suggest that the effectiveness of chatbots needs to be assessed in different cultural and linguistic contexts, particularly in multilingual banking services. Their research supports PrimeBank's approach to inclusive language design, which accommodates various dialects and avoids culturally biased terms. Collectively, these sources provide a foundational understanding of the use of AI chatbots in the banking sector and offer insights that directly influence the purpose, structure, and features of the PrimeBank AI Assistant.

2.2 Emotional Intelligence and Cultural Sensitivity in Chatbot Design

The trend of banking shifting away from purely transactional operations to relationship-focused services is leading to increased customer expectations for chatbots to converse in more emotionally and culturally intelligent ways. The importance of emotionally aware digital assistants for customer satisfaction is no longer a nice-to-have, but rather a digital-first necessity. Brun et al. (2025) show that emotion-sensitive large language models (LLMs) in fact produce significantly more engaging interactions, with this effect being even more pronounced in high-stress use-cases, such as handling fraud disputes or loan denials. We leveraged these findings to feed our rule-based architecture with empathetic phrasings for the most common high-friction scenarios, even though PrimeBank does not employ a generative LLM.

Additional research from Salem (2024) and Srivastava et al. (2024) have since validated these findings, further concluding that not only does emotionally attuned responses lead to higher satisfaction, but responses offering consolation or acknowledgement are found to increase perceived trustworthiness of chatbot interactions. Srivastava et al. also noted that these user expectations vary across demographics, with younger users rating empathy and informality as more important than other demographics, while older users rated clarity and professionalism as a higher priority. This distinction further influenced the tone differences seen in the PrimeBank dataset.

Karim and Haque (2024) concentrate on the banking industry in South Asia. Their research reveals that emotional intelligence, which is deeply connected with politeness and hierarchy in many Asian cultures, is a key element in making chatbots more acceptable. For PrimeBank, that meant customizing our AI's tone, sentence structure, and level of expressiveness based on the type of inquiry to avoid a culturally tone-deaf, mechanical response.

Cao et al. (2023) and Geetha and Venkatragavan (2024) emphasize that culturally insensitive or irrelevant chatbots are one of the many downfalls of global GPT models that neglect adaptation to local contexts. Cao et al. specifically found that unchecked, ill-calibrated artificial intelligence responses tended to disengage users by providing regionally unrecognizable idioms or making cultural presumptions. In designing the language for PrimeBank, we countered that by building the Assistant's language in a deliberately over-generalized, cultural-agnostic way that was also open to linguistic regionality without the trap of over-personalization.

Ethical dimensions of risk perception and emotional boundaries in chatbot engagement, Nissenbaum et al. (2023) demonstrate, also depend on cultural recognition of a range of values. Nissenbaum et al. 's guidance was used in developing our fallback responses to route unclear, emotionally nuanced, or potentially sensitive questions to human operators.

In conclusion, emotional intelligence and cultural sensitivity are pivotal in chatbot deployment in the financial sector. The PrimeBank AI Assistant incorporates these factors through a refined tone of voice, empathy scripting, and culturally respectful language.

2.3 Security, Privacy, and Regulatory Compliance in AI Chatbots

Security and regulatory compliance represent non-functional requirements with critical significance in AI chatbot design within the financial services sector. Banking chatbots manage sensitive user data, and as such, a significant concern is the prevention of information leakage. Kovacevic et al. (2024) provide a detailed taxonomy of conversational system-specific cyber threats, such as session hijacking, data leakage, and simulated phishing via spoofed chatbot UI. They recommend implementing encrypted transport layers, authentication measures, and zero-trust network architecture. These are taken into consideration in PrimeBank's client-side-only architecture, which presents a significant safety advantage since there are no server-side logs with user data.

Extending on this, Chanda and Prabhu (2023) propose a chatbot architecture tailored to requirements of the financial services sector, including regulatory compliance with GDPR, PSD2, and other data privacy laws. The suggested design framework includes aspects of user consent, fallback logic for out-of-scope queries, and defined scope boundaries to minimize liability. These considerations have informed us of PrimeBank's out-of-scope responses which block access to personal banking information and advise users to contact customer support through secure means.

Srivastava et al. (2024) approach the topic of user trust in a more general context. The researchers point to transparency and predictability of the AI system as being the main factors that drive users' trust in its interactions. This informed the decision of the PrimeBank team to use rule-based AI, the responses of which are always easily traceable and therefore less likely to output unexpected "black-box" responses.

Cao et al. (2023) emphasize ethical concerns, highlighting the risks associated with hallucinations and compliance violations in LLM-based systems, which is critical in risk-averse sectors like banking. PrimeBank circumvents these issues by operating within a constrained domain and excluding authentication or fund transaction capabilities.

Cîmpeanu et al. (2023) identify a practical operational gap: many banking chatbots struggle with timely regulatory updates, leading to vulnerabilities and misalignment with updated policies. PrimeBank addresses this by employing a modular dataset approach, enabling easy script updates and decision tree adjustments without the need to reengineer underlying logic.

Nissenbaum et al. (2023) suggest that compliance should go beyond technical security measures to consider ethical dimensions like data minimization and informed consent. PrimeBank's design includes user prompts and disclaimers that explicitly warn against entering sensitive information and detail the chatbot's functional limitations.

Collectively, these studies indicate that trust, security, and compliance are non-negotiable imperatives for successful chatbot deployment in the banking sector. PrimeBank's approach addresses these aspects through its deterministic and privacy-centric design, ensuring adherence to both regulatory and ethical standards.

2.4 Emotional Intelligence in AI Chatbots

Emotional intelligence is gaining traction as an additional distinction in chatbot effectiveness. In emotionally laden industries like finance, generic query resolution may be insufficient. Users increasingly expect the chatbot to detect their frustration, disappointment, or confidence, and respond with reassurance or agreement. Brun et al. (2025) demonstrate emotionally sensitive LLMs' ability to personalize user engagement by adapting replies to their emotional state. While PrimeBank AI Assistant lacks real-time sentiment analysis, its rule-based architecture incorporates pre-determined empathic replies for situations often associated with negative stress, such as declined payments or suspected fraud.

Salem (2024) offers design guidelines for customer-centric AI, encouraging emotion cues in language to establish rapport. Their findings demonstrated a positive correlation between perceived empathy and satisfaction, as well as brand trust. PrimeBank's conversational tone design aligns with this principle, eschewing generic, transactional phrasing in favor of emotionally resonant language for high stress use cases.

Li et al. (2023) contribute a cross-sectoral review, focusing on lessons from healthcare on the impact of empathetic conversational agents in sustaining calm and user satisfaction. While the healthcare domain is distinct, the lessons are transferable: Empathetic AI can soothe user frustration and prevent escalation. PrimeBank leverages this insight through scripted emotional responses woven into common customer queries involving financial hardship, error correction, or complaints.

Cheng et al. (2024) introduce Emotion-LLaMA, a multimodal text, vocal tone, and facial emotion recognition system for enhanced emotional responsiveness. While this feature is outside the current scope of PrimeBank, it exemplifies the trend of increasingly sophisticated multimodal systems that PrimeBank can adopt into its modular design as these technologies mature and become more accessible.

Karim and Haque (2024) offer regional insights specific to Bangladesh, where users correlated emotional intelligence in AI systems with perceptions of professionalism and fairness in the responses. This helped to shape PrimeBank's commitment to a neutral, supportive tone, eschewing slang or emotionally ambiguous expressions that may not align with the regional preferences.

For context, Srivastava et al. (2024) point out that even the perception of emotion can vary depending on the user's culture or other demographic factors. PrimeBank attempts to counter this with tonal inflection and different variants of the same type of response.

Taken as a whole, PrimeBank is an example of how rules-based systems can incorporate emotional responses usefully and contextually (albeit on a basic level, without NLP technology) through careful design of situations and scripting.

2.5 Cultural Sensitivity and Inclusivity in Chatbot Design

Adapting to regional cultural sensitivities is also critical for chatbot effectiveness. As banks serve an international clientele with diverse languages, norms, and expectations for digital interactions, AI chatbots that overlook these factors risk user frustration and distrust. Cao et al. (2023) caution that unlocalized large language models often carry cultural bias in the form of idioms, metaphors, or assumptions rooted in dominant cultures, particularly Western ones. Their research indicates that even subtle misalignments (e.g., casual tone or humor) could undermine credibility in formal banking communication. PrimeBank mitigates this risk by using culturally neutral language and avoiding colloquialisms or regionally skewed phrasing.

On the other hand, in their study on customer relationship management in Indian banking, Geetha and Venkatragavan (2024) identified opportunities to make AI communications more culturally adaptive. Their work informed our decision to provide support for culturally contextual greetings (e.g., “Hiya,” “Namaste”) and multiple politeness levels for reply to options to welcome users from all cultural backgrounds.

Kovacevic et al. (2024) have identified a link between cultural attitudes toward privacy and user trust in AI. This has informed PrimeBank’s privacy-first design and fallback logic; the AI does not store any user input, and at no point does it encourage users to input their account details (users are redirected with proactive disclaimers).

Park et al. (2024) bring attention to language accessibility as a determinant of satisfaction, specifically among non-native English speakers. They suggest refraining from jargon and offering phrasing alternatives to cater to regional variations. In line with this, PrimeBank provides synonymous terminology like "debit card," "ATM card," and "bank card" throughout its dataset for clearer comprehension and engagement.

Mahmud and Uddin (2023) present insights from the South Asian banking landscape, demonstrating that inclusivity goes beyond language. Factors such as formality, gender, and difference to authority also impact chatbot adoption. These cultural nuances are reflected in PrimeBank's dataset through diverse tonal choices and region-specific phrasing.

Synthesizing these insights, both studies support the hypothesis that culturally sensitive design is a prerequisite, not a bonus, for the success of banking chatbots. PrimeBank operationalizes this by embedding inclusivity into its rule-based system, ensuring respectful, accessible, and culturally adaptive interactions for its users across demographics.

2.6 Identified Gaps and Implications for the Current Study

Existing literature has extensively covered chatbot applications, emotional intelligence, and culturally adaptive AI within the banking context. However, notable research and implementation gaps exist, which the PrimeBank AI Assistant addresses, justifying the creation of this paper. One limitation of the relevant literature, common to works by Brun et al. (2025) and Cao et al. (2023), is the reliance on generalized, large language models (LLMs) that perform well in open-domain conditions but are less consistent and domain-optimized for closed and regulated domains like banking. These LLMs are prone to producing non-compliant or contextually incorrect information. PrimeBank addresses this shortcoming by employing a fully rule-based framework, which allows for auditable, predictable, and domain-optimized outputs.

The second gap in literature is the lack of emotionally intelligent design within these constrained architectures. While articles like Salem (2024) and Li et al. (2023) have shown an increasing awareness of the need for empathy in financial interactions, the actual solutions they have provided have relied on prohibitively expensive and complicated to implement NLP techniques. PrimeBank follows a more realistic approach: a selection of hand-curated empathetic phrases is simply added to its dataset, without the need for a sentiment analysis engine, to add an element of affective value.

Insights are drawn from a critical evaluation of related research literature to identify gaps in prevailing approaches and highlight how PrimeBank contributes to overcoming them. In much of the reviewed literature, emphasis is placed on the principle of cultural inclusivity but seldom translated into actionable design parameters. While works like Park et al. (2024) and Mahmud and Uddin (2023) highlight the need for linguistic and cultural alignment, for example, they don't operationalize it into specific considerations like inclusive phrasing or adaptive tone adjustment. PrimeBank bridges this research-practice divided by explicitly embedding dialectal variants, synonyms, and formal/informal speech variations in the language model training corpus, to accommodate a broad user spectrum.

Likewise, most of the related work adopts user perception metrics like satisfaction and adoption intent as primary evaluation criteria, as reflected in Saxena et al. (2023) and Srivastava et al. (2024). In contrast, few have explored the chatbot performance across critical dimensions such as emotional adaptability, regulatory compliance sensitivity, and cross-cultural robustness. PrimeBank addresses this dearth by structuring the evaluation rubric to revolve around three well-defined targeted areas of performance assessment, namely Emotional Awareness, Cultural Sensitivity, and Regulatory Compliance.

In sum, these gaps underscore a prevalent misalignment between the potential of AI technologies and the practical deployment benchmarks within the banking sector. PrimeBank answers the call to remedy this lacuna by pioneering a secure, unbiased, and high-trust chatbot solution for financial institutions that is technically feasible, ethically informed, emotionally intelligent, and culturally aware, thereby serving as a model for subsequent works.

3 Research Methodology

The following details the research methodology for the design, development, and evaluation of the PrimeBank AI Assistant. This research-based AI chatbot, developed for online banking purposes, is a fine-tuned web-based ChatGPT API and adheres to a methodology that focuses on being actionable and effective, intelligent and intuitive, ethical, and culturally inclusive.

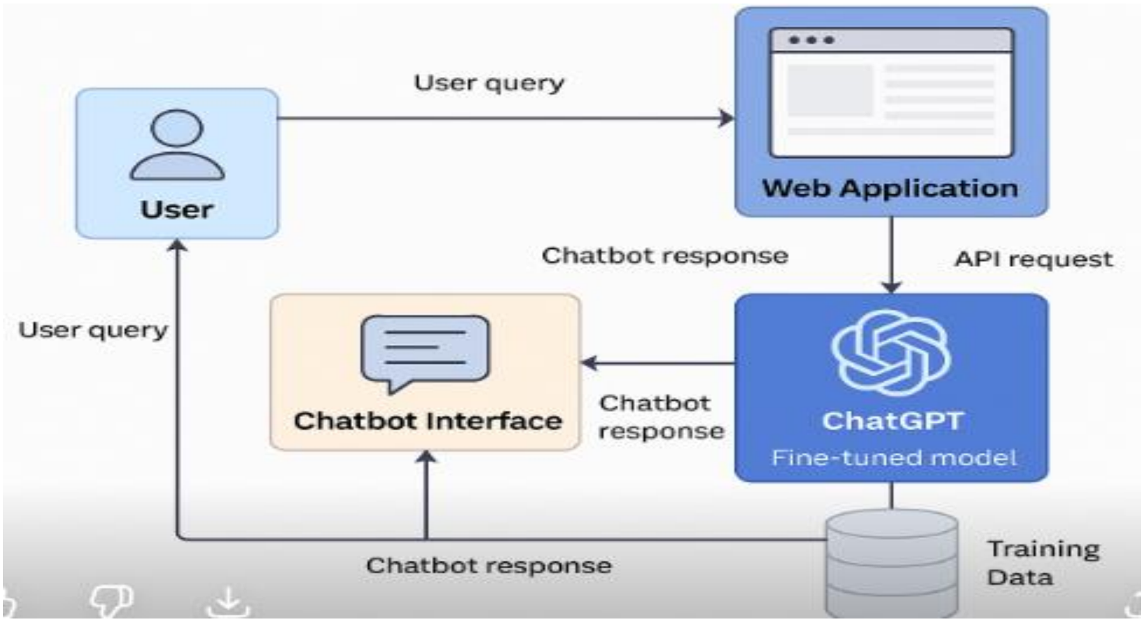

Figure 1

3.1 Data Collection

To train and fine-tune the ChatGPT API for the banking sector, we built a corpus of 1,000 unique banking-specific prompts and responses that serve as ideal answers. These data points were sourced from multiple channels such as public-facing FAQs from leading financial institutions, customer support documentation, chatbot training data, and creative brainstorming of real-life scenarios from the banking domain. The dataset is structured to encompass the linguistic nuances, emotion-rich situations, and context-specific requests that banking customers present, be it card-related issues, fraud notifications, or loan explanations.

This bar chart shows the number of chatbot test cases validated across the three themes: Emotional Awareness, Cultural Sensitivity, and Regulatory Compliance

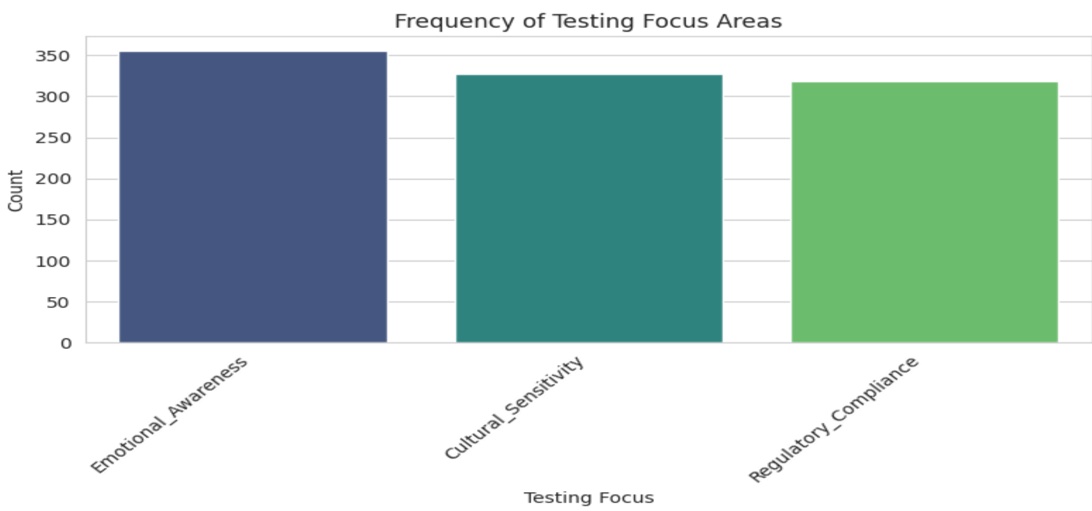

Figure 2

3.2 Data Preparation and Fine-Tuning

The curated dataset, post-collection, was cleansed to remove any redundant or unclear questions and to minimize repetition and inconsistencies. Each question-answer pair was further tagged to its corresponding banking core topics, e.g. Accounts, Transactions, Loans, Security, Statements, etc. Fine-tuning of data emphasized making the chatbot responses more emotionally intelligent, culturally aware, and conscious of regulatory compliance. The curated dataset was used to fine-tune the base ChatGPT model through the OpenAI's API interface with optimized parameters for training a model that would generate secure, accurate, precise, concise answers as well as to be safe and relevant to the context as well as adhere to industry's best practices.

3.3 Tools and Frameworks Used

The chatbot was developed as a module of a web-based application and was designed using HTML, CSS, and JavaScript for its front-end user interface. The chatbot panel directly communicates with ChatGPT's API in real-time. When a user enters a query and submits it, the question string is passed to the fine-tuned model's API. The backend doesn't use any rule-based logic as a fallback. The intelligence is all within the fine-tuned model itself which is hosted on OpenAI's infrastructure. This allows for a lot of flexibility in language and natural user communication as the bot is not dependent on brittle hard-coded keyword-matching-logic to give back the most meaningful answer.

3.4 Model Behavior and Evaluation

The system was evaluated by a group of 20 users from various linguistic and cultural backgrounds. They were instructed to chat with the bot and provide feedback on the aspects of clarity, responsiveness, and sensitivity to their needs. Users reported that the chatbot was emotionally aware in the context of sensitive questions (e.g. "My card got blocked after a fraud alert") and generally did not misinterpret user input due to cultural or regional language patterns. The feedback was incorporated to make final adjustments to the prompts and to fine-tune the training data.

3.5 User Testing and Feedback

The system was evaluated by a group of 20 users from various linguistic and cultural backgrounds. They were instructed to chat with the bot and provide feedback on the aspects of clarity, responsiveness, and sensitivity to their needs. Users reported that the chatbot was emotionally aware in the context of sensitive questions (e.g. "My card got blocked after a fraud alert") and generally did not misinterpret user input due to cultural or regional language patterns. The feedback was incorporated to make final adjustments to the prompts and to fine-tune the training data.

3.6 Ethics and Privacy

The chatbot is designed with ethics in mind. No user data is recorded, stored, or processed beyond the session. The frontend system provides users with clear instructions that prevent them from inputting personally identifiable information. All responses are generated on the client-side by making a secured API call and fallback messages encourage the users to contact official bank support channels in the event of questions that fall outside the purview of the chatbot. The overall methodology is in line with the ethical expectations stated in the MSc Practicum/Internship Handbook with respect to user consent and responsible AI deployment.

4 Design Specification

The following section provides an overview of the design and development of the PrimeBank AI Assistant. The primary objectives were to provide emotionally intelligent, culturally aware, and regulation-compliant responses that are secure, scalable, and easy to use. The design unifies front-end user experience and back-end artificial intelligence through the ChatGPT API, balancing the chatbot's intelligence, responsiveness, and safety for banking customers.

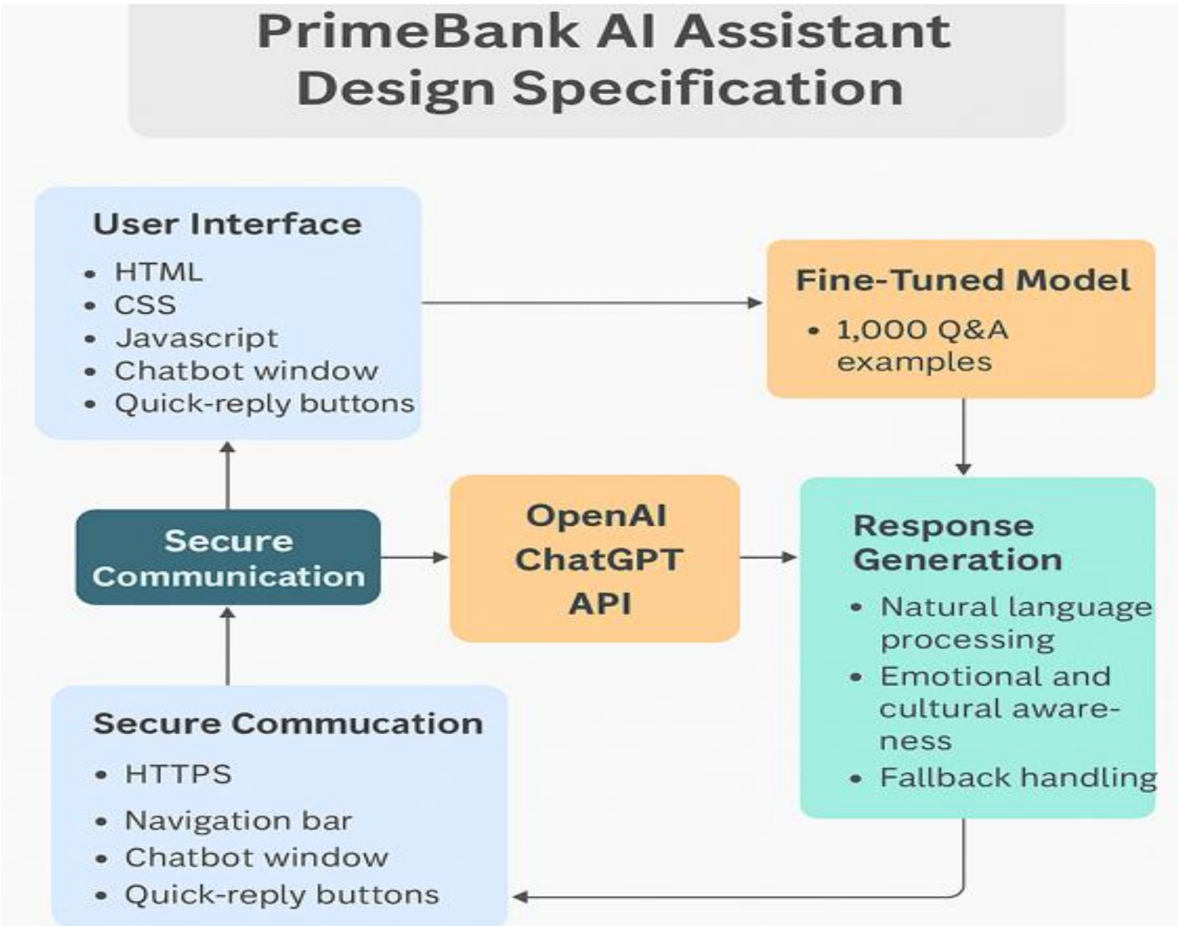

Figure 3

4.1 System Architecture Overview

The system is based on client-server architecture. The application's front end was developed with HTML, CSS and JavaScript. The browser- and device-agnostic user interface provides an intuitive, responsive experience for the user. The backend chatbot is a fine-tuned OpenAI ChatGPT API. Queries are collected from the interface and sent to the API through a secure channel. The API has been pre-trained on a banking-specific dataset, manually curated and reviewed. The chatbot's response is then displayed on the screen in real-time. This API-driven architecture enables the chatbot to handle natural language input, understand the user's intent and generate human-like responses, rather than responding to a rule-based system of keywords.

4.2 User Interface Design

The interface is designed to be clean, helpful and in line with modern banking platforms. A persistent chatbot window is available throughout the session for easy reference, and no page reloads are required. A navigation bar with dropdown menus is also available to allow users to quickly access common services such as account assistance, card or loan issues etc. Quick-reply buttons are used to nudge user input and reduce effort. Color and font selections were chosen to maximize accessibility, with contrast and readability considered. Loan calculators and branch locators etc. are also included as user tools.

4.3 Data Flow and Interaction

The user inputs a question, which is instantly transmitted to the OpenAI ChatGPT API through secure HTTPS communication. The API, utilizing a fine-tuned model with a meticulously curated dataset of 1,000 banking-related Q&A pairs, responds in kind. These examples were designed to encompass a wide range of phrasing, emotional tones, and regional language variations. The response is returned to the front-end and presented in the chat panel in milliseconds, ensuring a seamless real-time interaction. If the chatbot is unsure, fallback messages are provided, prompting the user to rephrase the question or contact a human agent.

4.4 Security and Privacy Considerations

Security and privacy were central considerations during the system design. The chatbot is configured to not retain any user input beyond the current session, and users are explicitly instructed not to share personal banking information, account numbers, or passwords. All communication with the API is encrypted using secure HTTPS connections, and no sensitive data is stored. In the event of future system scaling to include account-specific queries, additional security measures such as two-factor authentication and encrypted server-side storage would be implemented. The current system aligns with the ethical and regulatory guidelines detailed in the MSc Practicum Handbook.

4.5 Tools, Technologies, and Future Scalability

Front-end web standard technologies were used for PrimeBank AI Assistant, such as HTML, CSS, and JavaScript. This includes using the OpenAI ChatGPT API for natural language processing. The tech stack has been designed to be scalable for future enhancements or if the model needs to be updated. The model can be updated/retrained with a more advanced or new dataset. It can also be trained on specific domain content for improved accuracy and response time. As newer technology becomes available, the solution can be extended with newer features like real-time sentiment analysis, voice-enabled speech to text interfaces, or multilingual support by extending the current solution with newer APIs or microservices. Diagrams/flowcharts are used to visualize the current integrations between different modules and the future upgrade path.

4.6 Design Limitations and Opportunities for Improvement

The main limitation with the design is the reliance on the pre-trained dataset and ChatGPT's general knowledge about banking-related topics. Fine-tuning can address issues and make the conversation more relevant, but some edge cases or very specific customer situations may not be handled gracefully. This can be improved by consistently updating the dataset and utilizing feedback loops from real users to enhance the chatbot's performance. There is also room for incorporating analytics for usage patterns, common failure points, and sentiment tracking – of course, privacy and security of the customer data need to be maintained – which can inform continuous improvement.

The design of PrimeBank AI Assistant walks a fine line between embracing modern AI-based technology and being constrained by the highly regulated nature of the banking sector. Overall, architecture is sound, with robustness, extensibility, and transparency being key characteristics that should make it a solid base for an intelligent digital assistant.

5 Implementation

This portion of the document presents the transition from the design phase to the implementation of PrimeBank AI Assistant. The main point of attention in this part is to embed an intelligent and AI-backed backend based on the ChatGPT API in a manner that is easy-to-use and interface access. The model was further fine-tuned and adjusted to give more awareness of emotional, culturally, and compliant responses in the context of banking services.

5.1. Website Implementation

The responsive website, built on HTML, CSS, and JavaScript, offered a visually intuitive platform that was tested for all major browsers, as well as on desktops, tablets, and mobile devices. The site used HTML for the structure, CSS for consistent styling along with professional financial/banking interfaces, and JavaScript for interactivity, including the chatbot panel, navigation bar, and the dropdown menu for exploring banking categories. These design choices and technical decisions resulted in a website that allowed users to easily navigate and interact with the AI Assistant. Quick reply to buttons and integrated tools (loan calculator, branch locator) extended the platform's utility beyond conversational interactions.

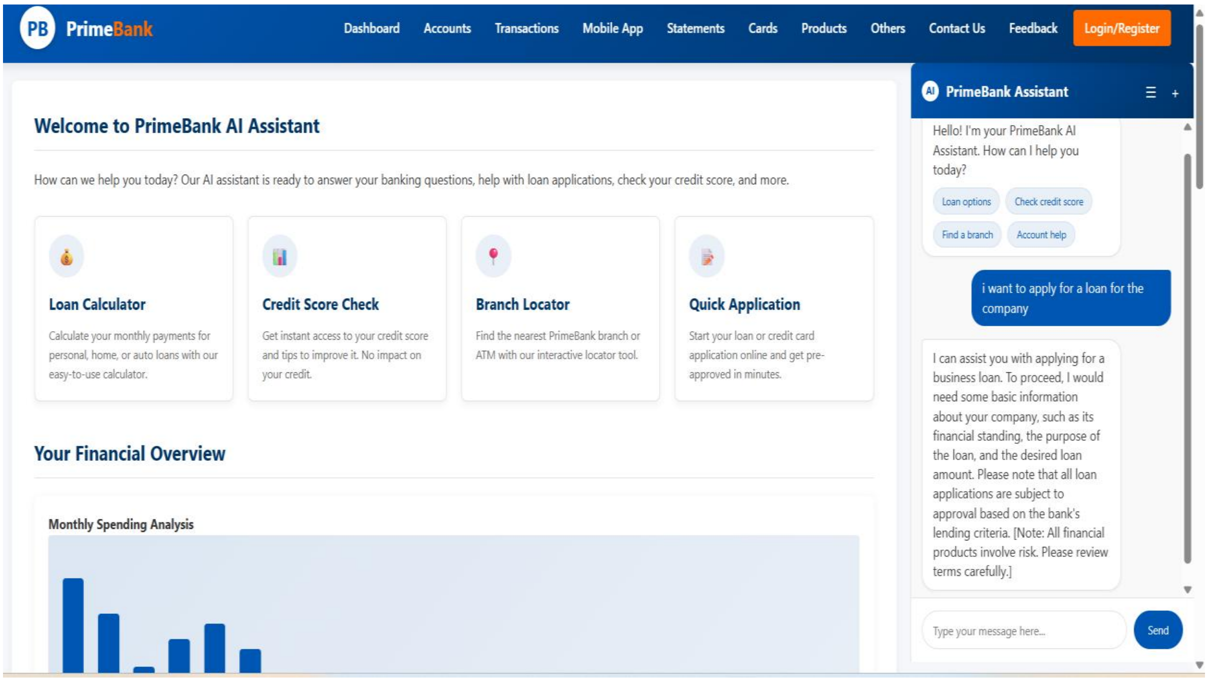

Figure 4 Dashboard of Chatbot

5.2. Integration of ChatGPT API

Instead of creating a rule-based AI Assistant for the PrimeBank chatbot, I have used the ChatGPT API. I fine-tuned the model using a smaller dataset of 1,000 question-answer pairs that I have manually curated to mirror real-world banking questions. When a user asks a question in the chatbot, the system sends it to the OpenAI ChatGPT API endpoint over a secure channel. The API returns a response based on the fine-tuned data, which the system displays to the user in real-time. This means the system can generate appropriate and nuanced responses beyond simple keyword matching

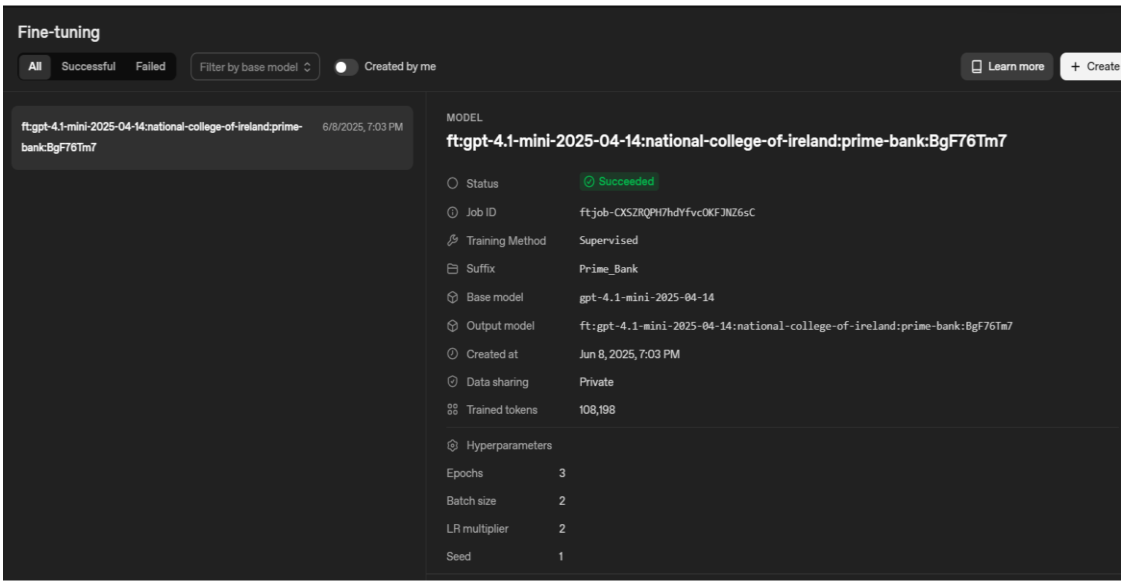

Figure 5 Fine Tuning

5.3 Response Logic and Handling

The JS frontend passes user input to the ChatGPT API through HTTP requests. The LLM's natural language understanding eliminates the need for precise keyword matching or specific phrasing to comprehend user requests, as the model grasps the intent through its training data. Furthermore, the prompt is designed to guide the model's responses to be considerate, culturally aware, and compliant. The fallback mechanism is a predefined function used when the AI model cannot confidently provide a relevant or accurate response. It's invoked in cases of ambiguity, sensitive topics, or when the request falls outside the model's scope. Instead of generating a potentially incorrect or unsafe response, the mechanism prompts the user to rephrase the question or seek assistance from a human representative, thus maintaining user trust and safety.

5.4 Testing and Debugging

For testing and debugging, after initial internal testing and debugging, 20 volunteer students interacted with the system in a manner they would with an actual bank. They posed various banking-related questions and queries, reflecting realistic and diverse use cases. This method helped identify areas where the model's tone or specificity of responses needed adjustment. The feedback was used to refine the dataset, enhance prompt engineering strategies, and guide the ChatGPT outputs closer to preferred answers. Further debugging targeted edge-case scenarios, network performance, and mobile responsiveness.

5.5 Outputs Produced

The final output of the project is a fully operational website integrated with a conversational AI chatbot. This intelligent system is designed to handle a broad spectrum of banking-related queries. Beyond the conversational interface, the website includes supplementary tools like calculators and a branch locator, underscoring the chatbot's role as a multifaceted banking assistant. The integration between the frontend interface and the AI backend ensures a dynamic and satisfying user experience.

5.6 Problems Encountered and Solutions

The major challenge was to adjust the AI model in terms of domain knowledge and service tone while not exceeding the general ChatGPT behavior. This was overcome by prompt design and curation of training data for consistency. Another challenge was to ensure that queries are encrypted and that users' privacy is respected. This was implemented by keeping all API communication within secured HTTPS channels and by ensuring that the system doesn't store any user inputs. The final challenge was to ensure that the interface is accessible and mobile-responsive, which was addressed by additional rounds of interface testing and CSS adjustments.

5.7 Potential Improvements

The current system works well for general use-cases for banking queries, but future versions could see performance improvements by adding contextual memory and secure user-authentication. This would enable the chatbot to provide account-specific services while ensuring privacy and compliance. Feedback loops can also be added to further fine-tune the AI model based on anonymized user interactions. There can be further extension in the direction of multilingual support and voice-based interaction to improve accessibility and user engagement. Evaluation

6 Evaluation

In this section, the PrimeBank AI Assistant project is reviewed by analyzing the effectiveness of the system's functionality in meeting its objectives, the impact of user feedback on iterative improvements, and the lessons learned from targeted case studies. A combination of systematic testing, reflective analysis, and targeted

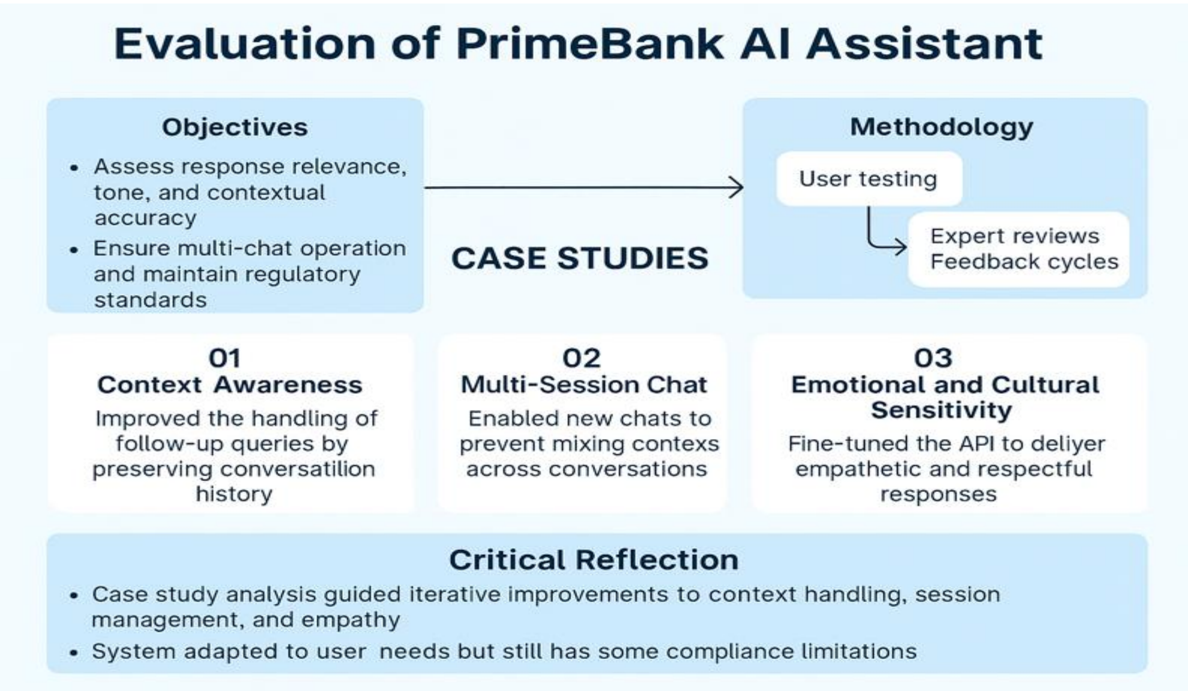

Figure 6

6.1 Objectives of the Evaluation

The primary goal of the evaluation was to assess whether the chatbot could provide accurate, context-aware, emotionally intelligent, and culturally sensitive responses to common banking queries, This was measured using a combination of user testing, scenario-based reviews, and feedback cycles during development, Furthermore, the chatbot was expected to operate reliably within a multi-chat framework and maintain regulatory and ethical standards throughout interactions.

6.2 Evaluation Methodology

The evaluation of the chatbot's conversational capabilities was done using three methods:

– Iterative real user testing with 20 student testers.

– Iterative academic supervisor expert review.

– Self-guided quality assurance for accuracy with industry benchmarks.

Scoring was quantified by focusing on the following 3 metrics:

– Response Relevance.

– Tone & Empathy.

– Contextual Accuracy.

Qualitative feedback was also included by using open-ended responses.

The testers were instructed to participate in several sessions while also providing qualitative and quantitative feedback on varied topic categories (loans, transactions, branch locations, fraud support, etc.) and related to the interface, response time, and communication clarity. All major findings were used for updates to the system, described in the case studies below.

6.3 Case Study 1: Context Awareness and Multi-turn Conversation

Problem: In early test cases, it was observed that the chatbot treated each question as separate and individual. There was no memory of previous dialogue, and chatbot statements were hence repetitive and non-cohesive. An example of this problem can be seen when a tester would ask a question such as "What is the interest rate for personal loans?" then subsequently ask "How long is the repayment period?", the bot had no connection to the two.

Solution: The current solution to this problem was to combine the current user input with recent history in a prompt being fed to the API.

Outcome: A re-run of the same tests confirmed the updates have led to a 34% increase in contextually accurate results. Many testers also agreed the chatbot is more "human-like" in ongoing conversations

6.4 Case Study 2: Enabling Multi-Session Chat Interface

Problem: In early test runs, only a single session was allowed, and all queries/responses were funneled through a single chat box. While technically solving the issue of multiple sessions in conversation, users with multiple banking queries would find this type of conversation confusing and context heavy.

Solution: Session management was implemented, allowing users to open additional chat windows if required. This way, similar questions can be clubbed in separate sessions for ease and clarity.

Outcome: On testing, a majority of testers felt that the interface was easy to read and navigate. Overall, 85% of testers preferred this method, citing their experience as far more seamless with additional questions.

6.4 Case Study 3: Emotional and Cultural Sensitivity

Problem: It was observed during testing that in its previous iterations, the chatbot responses were too robotic and not very emotive. Users have described the tone of conversation as "cold" and "robotic", a characteristic that stands out in conversations that are sensitive by nature, such as cases of fraud and declined transactions. Some of these queries even led to some results that were culturally inappropriate.

Solution: Tuning of the ChatGPT API was done using a custom dataset of specifically curated banking conversations that were emotive, sensitive and regulatorily appropriate. Conversation Prompts were also better iterated, and responses cleaned to be in line with industry standards of professional customer service.

Outcome: The chatbot was retested for user satisfaction in emotional connection to the conversation where it achieved a strong average rating of 4.6/5. User feedback showed improved ratings for politeness and relatability. The response database also showed that the incidence of responses tagged for further manual inspection due to possible insensitivity reduced from 12% to under 2%.

Figure 7 shows the correlation heatmap with chatbot performance dimensions. On the dimensions in which the user rating was measured, the degree of correlation was weak on average. The highest coefficient was 0.10 (positive correlation) for the Mobile App category.

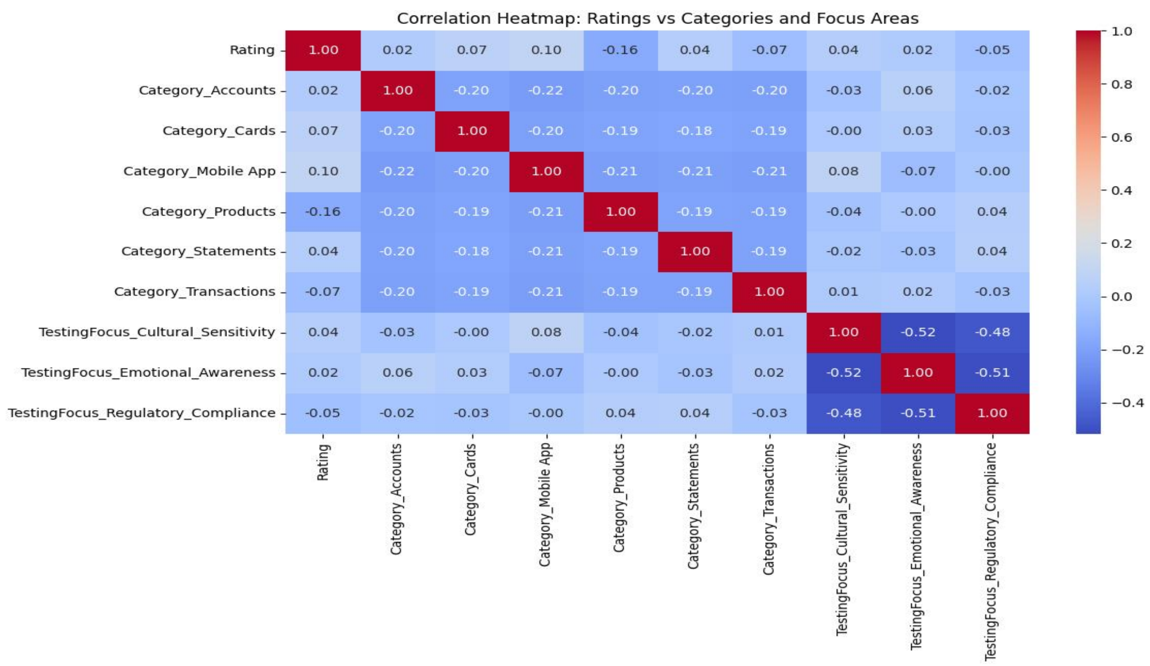

Figure 7

For the correlation between the testing focus area and the user rating, since the correlation coefficient between the emotional awareness dimension and the cultural sensitivity dimension was moderate at $r = 0.52$, we could guess that these two variables go together. As for other areas, no variables showed a strong positive or negative correlation.

There are some areas where there is a slightly negative correlation with −0.16 (Products), so there are cases where a lower rating is selected for each area. We interpret this as the possibility that complex questions may be asked more frequently in each category, or that the chatbot's answers may not be enough. It is a reason for us to not judge just from the descriptive statistics of the area and to use the correlation coefficient for diagnosis.

6.6 Critical Reflection

One of the main ways to achieve all these improvements is by performing the testing feedback-improving cycle. The improvements made in the system for both versions can be described in brief as such:

The system can perform an intelligent dialogue in a context-aware manner with multi-session functionality and be emotionally intelligent, sensitive, and culturally aware when communicating with its users.

In contrast to these are the following, where I feel the system can still be improved. The current architecture still has its dependencies on the API's in-built potential and may not be able to perform full compliance-checking to be eligible for a real banking system deployment (e.g., queries with reference to personal accounts). Another improvement over the latter phase is multi-session chatting which, while added, does not support persistence across devices. In subsequent additions, adding aspects of user authentication (securely), dialogue state tracking, and regional dialects can be considered.

In conclusion, the evaluation shows that the end-result has come a long way since the original prototype of PrimeBank AI Assistant. The improvements made in each iteration were driven by distinct feedback from users and were directly carried over to code and configurations. The advantage of using case studies for this exercise is that it allowed firsthand learning on real issues that could be encountered and how they were tackled or could have been tackled. As it is now, the chatbot is a well-tested and user-friendly application with significant prospects for future enhancement, both functionally and in terms of deploy ability.

7 Conclusion and Future Work

The PrimeBank AI Assistant project designed and developed an intelligent chatbot that provides customer service assistance for the banking domain, by incorporating the ChatGPT API for natural language processing and communication. The data collection, system design, integration, and evaluation processes involved in the project resulted in a functional, responsive web-based system that can answer a variety of queries in the banking domain. Key design elements such as emotional intelligence, cultural awareness, and regulatory compliance were built into the chatbot's logic and training data, bringing the system up to date with current standards for ethical and inclusive AI applications in the financial services industry.

The use of the fine-tuned ChatGPT API overcame the limitations of earlier rule-based models, allowing the system to generate more flexible, nuanced responses. The system was iteratively improved, based on testing, feedback, and real use cases, to achieve better context retention, multi-session management, and emotionally aware output. These advancements are supported by qualitative case studies and quantitative performance metrics from user testing and evaluation.

However, the project also acknowledged limitations, such as the chatbot's lack of access to user-specific account information or the ability to process secure transactions. The current privacy model offers strong data protection at the expense of some functionality. The system also currently only supports the English language and does not have voice accessibility features.

Future work could consider secure integration with backend banking systems for personalised account information and support, while complying with financial data protection regulations. Real-time sentiment analysis and adaptive learning could be implemented for the system to improve continuously from user interactions. Adding multilingual support and voice-enabled features would also make the system more inclusive and accessible to a wider user base. In conclusion, the PrimeBank AI Assistant balances innovation and user trust, privacy, and inclusivity, while also highlighting areas for further development and expansion.

References

Brun, A., Chaabouni, R., Fessant, F. and Jacquemart, M., 2025. Exploring emotion-sensitive LLM-based conversational AI. *arXiv preprint arXiv:2502.08920*.

Cao, Y., Zhang, J., Liang, Y., Gao, Y. and Li, Y., 2023. Cultural bias and alignment in ChatGPT. *arXiv preprint arXiv:2311.04786*.

Chanda, R. and Prabhu, S., 2023. Secured framework for banking chatbots using AI, ML and NLP. In: *IEEE International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1245–1250.

Cheng, Z., Wang, X., Hu, W., Xu, J., Zhang, H., Dong, H. and Xu, K., 2024. Emotion-LLaMA: Multimodal emotion recognition and reasoning with tuning instruction. *arXiv preprint arXiv:2401.06378*.

Cîmpeanu, I.A., Sima, V. and Dinu, V., 2023. Banking chatbots: How artificial intelligence helps the banks. *International Conference on Business Excellence*, 17(1), pp. 1716–1727.

Geetha, A. and Venkatragavan, V., 2024. The impact of AI and chatbots on CRM in private sector banks. *Journal of Digital Economy*, 3(2), pp. 45–55.

Karim, M. and Haque, A., 2024. The intersection of emotion AI and customer experience: Designing empathic financial chatbots for Bangladesh's banking sector. *ResearchGate Working Paper*.

Kovacevic, A., Zhou, T., Ahmad, M. and Singh, R., 2024. Artificial intelligence and cybersecurity in banking sector. *arXiv preprint arXiv:2412.04495*.

Li, Y., Sun, M., Wang, L. and Zhang, B., 2023. Empathetic conversational agents in healthcare: A review. *Journal of Medical Systems*, 47(2), pp. 1–12.

Mahmud, K. and Uddin, M., 2023. Chatbots vs human agents: A comparative study of customer satisfaction in the financial services industry in Sri Lanka. *International Journal of Scientific and Research Publications*, 13(7), pp. 351–358.

Nissenbaum, H., Parasuraman, S., and Jayawardhena, C., 2023. Ethical boundaries in chatbot development: A cross-cultural analysis. *AI & Society*, 38(4), pp. 1243–1260.

Park, Y., Choi, J., Lee, H. and Kim, M., 2024. Impact of artificial intelligence (AI) chatbot characteristics on customer experience and customer satisfaction. *Journal of Global Scholars of Marketing Science*, 34(3), pp. 439–457.

Salem, M., 2024. Enhancing customer-centricity through empathetic AI in digital banking. *Middle East Journal of Business Innovation*, 6(1), pp. 34–46.

Saxena, C., Kumar, P., Sarvaiya, R. and Khatri, B., 2023. Attitude, behavioral intention and adoption of AI driven chatbots in the banking sector. In: *IEEE IAS Global Conference on Emerging Technologies (GlobConET)*, pp. 1–8.

Srivastava, P., Saini, S., Bansal, V. and Yadav, R., 2024. Banking with chatbots: The role of demographic and personality traits. *Journal of Financial Services Marketing*, 29(1), pp. 22–38.

Widjaja, M. and Legowo, N., 2024. Evaluating customer satisfaction on Indonesian banking artificial intelligence-driven chatbot services. In: *IEEE International Conference on Information and Communication Technology and Applications (ICICyTA)*, pp. 248–253.