INTERNATIONAL JOURNAL OF SOCIAL SCIENCE, EDUCATIONAL, ECONOMICS, AGRICULTURE RESEARCH AND TECHNOLOGY (IJSET) | ISSN (E): 2827-766X
Website: https://ijset.org/ | Volumes 5 No. 3 (2026)

DESIGN AND DEVELOPMENT OF A GEMINI 2.0 FLASH-BASED CHATBOT FOR ANALYZING TRANSACTION DATA OF GRAB MSME PARTNERS

Lianawati Suwito[1]*, Sayekti Harits Suryawan[2], Hendra Saputra[3]
Program Studi Teknik Informatika, Fakultas Sains Dan Teknologi,
Universitas Muhammadiyah Kalimantan Timur[1,2,3]
E-mail: 2211102441024@umkt.ac.id[1]*, shs500@umkt.ac.id[2], hs048@umkt.ac.id[3]

Received:01 /01/2026 | **Revised**: 20/01/2026 | **Accepted:** 10 /02/2026 | **Published**: 28/02/2026

Abstract
The development of the digital economy requires Micro, Small, and Medium Enterprises (MSMEs) to utilize transaction data as a basis for business decision-making. However, limited digital literacy and technical capabilities have prevented many MSMEs from optimizing the use of data. This study aims to develop an Artificial Intelligence (AI)-based chatbot system capable of processing static transaction data sourced from the University of Malaya Malaysia Hackathon 2025 and providing simulated business insights and recommendations. The research method includes data preprocessing, rule-based intent detection implementation, and the integration of a Large Language Model (LLM) through the Gemini API to generate contextual responses. System evaluation was conducted using accuracy, precision, recall, and F1-score metrics, focusing on the chatbot's technical performance and its ability to produce consistent responses. The results indicate that the chatbot is capable of presenting information such as best-selling products, busiest ordering hours, and sales summaries, with stable technical performance. Therefore, this system demonstrates potential as an AI-based solution to assist MSMEs in understanding transaction data and supporting simulated data-driven decision-making.

Keywords: Chatbot, Artificial Intelligence, Large Language Model, CSV, MSMEs, Business Insights

INTRODUCTION

The digital economy in Southeast Asia has experienced rapid growth in recent years, driven by the emergence of multifunctional application services such as Grab, which provides online transportation, food delivery, digital payments, and integrated logistics services (Kee et al., 2021). The *e-Conomy SEA 2024* report projects that Indonesia's digital economy recognized as the largest digital market in Southeast Asia grew by 13% compared to the previous year, with a Gross Merchandise Value (GMV) reaching USD 90 billion in 2024 (Pichai & Hassabis, 2023). Micro, Small, and Medium Enterprises (MSMEs) represent both a sector significantly impacted by and a key driver of the economy that requires digital transformation, including through partnerships with platforms such as Grab (Latifah et al., 2026). Active MSME participation in this digital ecosystem creates substantial opportunities for leveraging data as a foundation for more effective business decision-making (Godwin et al., 2024).

However, many small business actors still rely on intuition and subjective experience in making business decisions, resulting in suboptimal use of available objective data (Katti & Mutmainah, 2020). Low levels of digital literacy and limited technical capabilities hinder entrepreneurs from utilizing technology to support business activities, including digital marketing and data-driven strategic decision-making (Sutisna et al., 2025). In this context, Artificial Intelligence (AI) technology offers a potential solution to help MSMEs improve operational efficiency and customer service (Fahmi, 2024). The use of AI-based chatbots, as one application of AI, has been proven to provide interactive, responsive, fast, and accurate services (Alsadoun & Alnasser, 2025). Although Natural Language Processing (NLP) models such as Large Language Models (LLMs) have begun to be adopted for interaction automation, their implementation still faces significant challenges in languages with limited linguistic resources, such as Bahasa Indonesia (Maria, 2024). Previous studies indicate that AI adoption among MSMEs has the potential to enhance operational efficiency, customer satisfaction, and competitiveness through automation and data analysis (Harahap et al., 2025). One effective form of implementation is the use of AI-based chatbots, which play a significant role in improving service efficiency and customer experience across various business sectors (Harisi & Hiwono, 2024a). Chatbot implementation on marketplace platforms has even been shown to reduce costs and time while improving user experience (Fitriani et al., 2022). Meanwhile, the use of LLM technology has demonstrated high effectiveness in various financial tasks, such as data analysis and reporting (Rane et al., 2024). Large Language Models (LLMs) such as ChatGPT and Gemini have advantages over smaller models because they can understand complex contexts and generate more natural text for tasks such as summarization, sentiment analysis, and data-based reporting (Minaee et al., 2025). Recent studies state that Gemini has advantages in factual accuracy compared to ChatGPT through its integration with Google Search, making it more suitable for developing data-driven service chatbots (Rane et al., 2024). Nevertheless, many MSME actors still determine products based on intuition without relying on historical data, making data-based product recommendations necessary to identify sales trends, high-demand products, and profit opportunities (Sutisna & Raudhan, 2025). Utilizing transaction data in the form of CSV datasets enables systems to provide product recommendations based on actual sales patterns in a simulated manner, making business decisions more objective and responsive to market dynamics (Festiyed et al., 2024). In addition, the utilization of AI technologies such as the Gemini API a recent technology from Google DeepMind remains very limited in the Southeast Asian context (Akhtar, 2024). Technological infrastructure and human resource readiness are also barriers to implementing AI solutions tailored to local needs (Anggraini et al., 2025). Therefore, there is an urgent need to develop an AI-based chatbot that is not only capable of presenting data concisely and informatively but also uses simple and easily understandable language for non-technical users, thereby improving digital literacy and empowering MSME actors to make data-driven business decisions independently.

This study aims to develop an AI-based chatbot system capable of automatically generating insights from MSME transaction data in multiple languages using technology from the Gemini API. The system is designed to recognize user intent so that the chatbot can provide accurate and relevant responses, such as information about "best-selling products" or "recommendations for new products to sell." With this solution, MSME actors, particularly in Indonesia, are expected to more easily understand and utilize their business transaction data without requiring specialized technical expertise (Suryanto et al., 2023). The implementation of this system not only provides practical benefits in improving operational efficiency and customer service but also contributes to the development of data literacy among MSMEs, which is an essential foundation for inclusive and sustainable digital economic development in Southeast Asia (Chitturu et al., 2017; Fahmi, 2024). The data used in this study were obtained from Grab merchant partners in Southeast Asia through the Grab Malaysia Hackathon 2025 at the University of Malaya, allowing the dataset to be used to evaluate the chatbot system's technical performance in a simulated environment.

LITERATURE REVIEW

Micro, Small, and Medium Enterprises (MSMEs)

MSMEs are commercial businesses managed by individuals that meet the criteria of productive economic enterprises as regulated under Law of the Republic of Indonesia No. 20 of 2008 concerning MSMEs (Utami & Sasmita, 2022). MSMEs constitute a primary pillar of the global economy due to their dominant number and significant contribution to job creation (World Bank, 2021). Digital transformation provides substantial opportunities for MSMEs to adapt to technology-based trading systems.

Gemini API

The Gemini API is an application programming interface developed by Google DeepMind that provides access to the Gemini model, one of the latest generations of Large Language Models (LLMs). LLMs such as Gemini are capable of performing various tasks, including automated text generation, conversational agents, and large-scale contextual analysis. The role of the Gemini API in research and industry is to provide a flexible technical foundation for developers to create intelligent and adaptive AI-based solutions. This demonstrates the close relationship between Artificial Intelligence (AI), Deep Learning, and Natural Language Processing (NLP), where Gemini integrates these approaches to achieve improved performance compared to previous generations (Pichai & Hassabis, 2023).

Artificial Intelligence

Artificial Intelligence (AI) is a field of computer science focused on developing systems capable of mimicking human intelligence, including the ability to learn, reason, understand natural language, and make decisions (Russell et al., 1995). One of the main branches of AI driving many current innovations is Deep Learning. This technology utilizes Neural Network architectures to recognize complex patterns in large-scale data, thereby supporting decision-making processes while reducing dependence on human intervention (Mijwil et al., 2022). The integration of AI into platforms such as GrabMerchant leverages these techniques to provide a more personalized, responsive, and efficient user experience. Modern AI development is also closely associated with the emergence of Large Language Models (LLMs), which increasingly dominate human language processing and understanding.

Natural Language Processing (NLP)

Natural Language Processing (NLP) is a branch of AI focused on enabling computers to process and understand human language (Jurafsky & Martin, 2000). NLP allows machines to recognize user intent (intent detection), identify important entities (named entity recognition), and generate relevant automated responses. Recent advancements show that NLP is increasingly supported by Transformer-based architectures, which form the foundation of Large Language Models (LLMs) such as ChatGPT and Gemini.

These two models possess different strengths: ChatGPT excels in conversational fluency and linguistic creativity, while Gemini demonstrates stronger factual accuracy and integration with external infrastructures such as Google Search (Rane et al., 2024). By leveraging Deep Learning techniques within the Transformer architecture, NLP can better understand conversational context and produce increasingly natural human–machine interactions.

Large Language Model (LLM)

A Large Language Model (LLM) is a type of AI model designed to process and generate textual language. LLMs are generally trained on massive amounts of textual data and employ Deep Learning techniques to learn language patterns and structures (Hadi et al., 2023). The primary method underlying LLMs is the Transformer architecture, one of whose key components is self-attention. Self-attention explicitly models interactions among elements within a sequence. This mechanism enables each element in the sequence to be updated by integrating global information from the entire input (Khan et al., 2022).

Chatbot

A chatbot is an AI-based program designed to conduct automated conversations with users, either through text or voice (Ardiansyah & Sulaksono, 2023). Chatbots function by integrating NLP and LLM technologies to:

- Understand user questions (Intent Detection).
- Retrieve relevant information (entity recognition and retrieval).
- Provide automated and contextual responses.

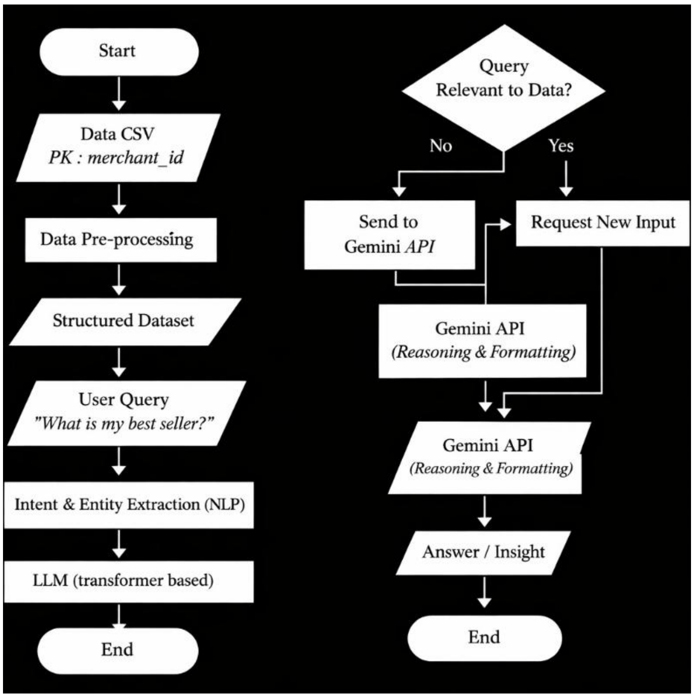
In the context of this research, the chatbot acts as a digital assistant for MSMEs, helping to provide product information, respond to customer inquiries, and support business management. By utilizing LLMs as a foundation, the chatbot can generate interactions that are more natural, accurate, and adaptive to user needs.

METHOD

This research focuses on Micro, Small, and Medium Enterprises (MSMEs), considering that approximately 90% of all companies in Indonesia are MSMEs, and they contribute to more than 90% of total employment absorption in the country (Novitasari, 2022). An AI-based chatbot was selected because of its ability to understand user queries and automatically provide data-driven responses. The utilization of Natural Language Processing (NLP) technology and Large Language Models (LLMs) such as Google Gemini enables the chatbot not only to respond to simple questions but also to deliver business analysis and insights in language that is easy to understand (Harisi & Hiwono, 2024b; Minaee et al., 2025). Thus, this study is directed toward developing a chatbot capable of providing business analysis support for MSMEs. The presence of the chatbot is intended not only to facilitate access to sales information and consumer trends but also to encourage MSMEs to utilize digital technology more optimally in order to enhance their competitiveness in the digital economy era.

The methodological framework is a structured sequence of steps used to implement techniques or approaches in achieving the research objectives, ensuring that each process is carried out systematically and in a well-planned manner. In the study titled *Development of a Simulation System for an MSME Business Insight Chatbot Based on Transaction Data*, the methodological framework illustrates the process of developing a data-driven chatbot system from initial data processing to generating responses or business insights for users. The complete methodological flow diagram is presented in Figure 1.

Figure 1. Methodological Flow of MSME Business Insight Chatbot Development



Description:

1. **Start**
   The process begins when the system is ready to receive merchant data and user queries.
2. **Merchant Data Input (CSV)**
   At this stage, the system receives merchant data in CSV format. The data contains key information such as product lists, prices, and menu descriptions, linked through a primary key in the form of merchant_id. This dataset serves as the system's knowledge base before the chatbot can provide recommendations or answer user questions.
3. **Data Pre-Processing**
   The received data may not be immediately ready for use. Therefore, a pre-processing stage is carried out to improve data quality. This process includes data cleaning (removing duplicates and null values), data transformation (text normalization and creation of new columns), and extraction of important features (menu, price, quantity, sales). The final result of pre-processing is a more concise and structured dataset ready for use
4. **Structured Dataset**
   The result of pre-processing is a well-organized and structured CSV dataset. With this dataset, the system can match user queries more quickly and accurately, ensuring that the responses provided are more relevant.
5. **User Query**
   The user submits a question, for example: "What is my best seller menu?" or other questions related to merchant data. This question serves as the entry point for the system to begin processing.
6. **Intent & Entity Extraction (NLP)**
   The incoming query is analyzed using NLP methods to understand the user's intent. This analysis includes:
   a. **Intent Detection** to determine the main objective of the question, such as identifying the best-selling menu or asking about prices.
   b. **Entity Extraction** to capture important details, such as menu names or specific categories mentioned in the query.

7. **LLM (Transformer-based)**
   The LLM is used to deeply understand the context of the question and connect it with the structured merchant data. Its role includes:
   a. Interpreting the user's query semantically.
   b. Constructing appropriate queries based on the merchant dataset.
   c. Serving as a bridge between natural user language and more technical data formats.
8. **Decision: Is the Query Relevant to the Data?**
   After identifying intent and entities, the system checks whether the question matches the merchant dataset.
   a. If relevant, the process proceeds to the Gemini API call stage.
   b. If not relevant, the system provides feedback to the user to revise or restate the question in a clearer format. However, even if the query does not match the dataset, the system will still generate a follow-up response.
9. **Send to Gemini API**
   Queries relevant to the data are sent to the Gemini API. This stage supports the reasoning process and prepares the response in a more natural format.
10. **API (Reasoning & Formatting)**
    The Gemini API processes the query by combining available data and generating a response that is easier to understand. The analysis result is not merely raw data but is formatted into a contextual explanation.
11. **Answer / Insight**
    At this stage, the system delivers the final response to the user. The answer may include relevant business insights, such as a list of best-selling menus along with additional details that enrich the information.
12. **Evaluation**
    The chatbot is evaluated using a confusion matrix by comparing the chatbot's responses with the ground truth from the merchant dataset. Testing uses a set of evaluation questions aligned with the chatbot's core functions. Performance metrics such as Accuracy, Precision, Recall, and F1-score are calculated to assess the system's accuracy and consistency.
13. **End**
    After the response is delivered, the interaction is considered complete. The system then returns to its initial state, ready to receive the next user query.

RESULTS AND DISCUSSION

Program Implementation Results

The successfully developed chatbot has been deployed in the form of an HTML-based website, and its interface can be seen in Figure 2.

Chatbot Interface Display Results

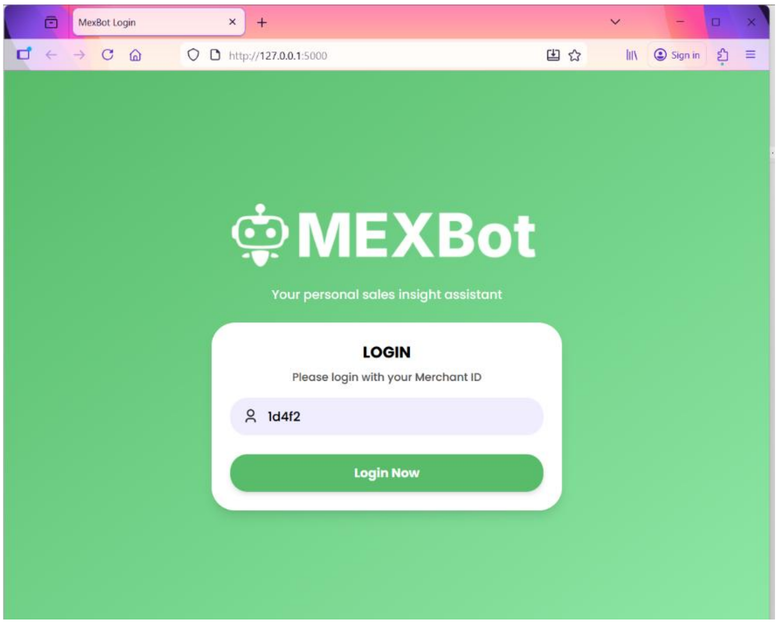


Figure 2. Chatbot Interface Display Results

Figure 2 shows the chatbot interface implemented as a web-based application. On the initial page, users are required to enter their respective merchant_id as a simple authentication mechanism and as an identifier of the transaction data source to be processed by the system. After successful authentication, the main chatbot interface is displayed, and users can begin interacting with the system through the available conversation input field. Functionally, this chatbot system is designed to use English as the primary language for presenting technical responses and insights, as the transaction data and the underlying model are oriented toward English-language datasets. However, the system also provides Indonesian language support as an additional feature to assist local users who may have limitations in understanding technical terminology in English. This bilingual feature is expected to enhance user experience, particularly for MSME actors with low to moderate levels of digital literacy. Furthermore, the chatbot interface is designed to be simple, lightweight, and responsive, allowing it to be accessed through various devices such as laptops and smartphones. Figure illustrates the initial layout of the chatbot, while Figures present examples of user interaction with the system and the resulting sales insights. This visual documentation provides a clear overview of how the chatbot operates within a web-based implementation context.

Responses to Greetings, Menu Requests, and Best Seller Queries

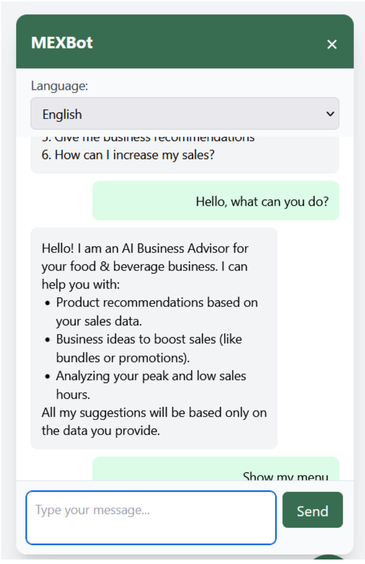


Figure 3. Responses to Greetings, Menu, and Best Seller

Figure 3 presents the chatbot's responses to user queries in the form of greetings, menu requests, and best-selling product inquiries. The chatbot successfully accesses and displays menu data as well as best-selling items sourced directly from the CSV file. As a result, the information provided is factual and data-driven, minimizing the risk of hallucination and ensuring response accuracy.

Results of Responses for Ordering Hours and Monthly Sales

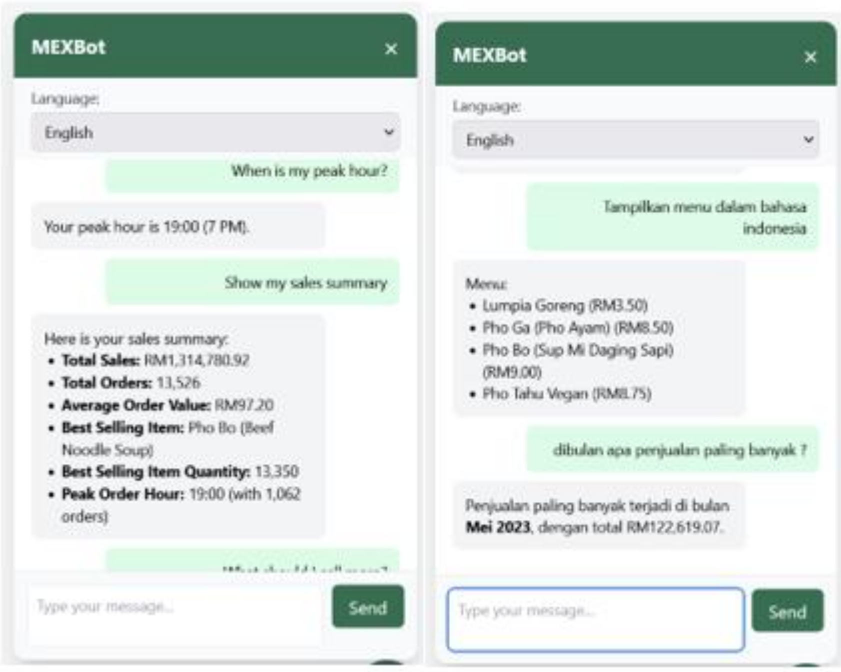


Figure 4. Responses for Ordering Hours and Monthly Sales

Figure 4 demonstrates that the chatbot is capable of providing structured summaries of data related to peak and lowest ordering hours, as well as monthly sales history. The generated responses are accurate, follow a chronological order (by date and month), and are entirely derived from the CSV dataset.

Data-Driven Product Recommendations and Business Ideas

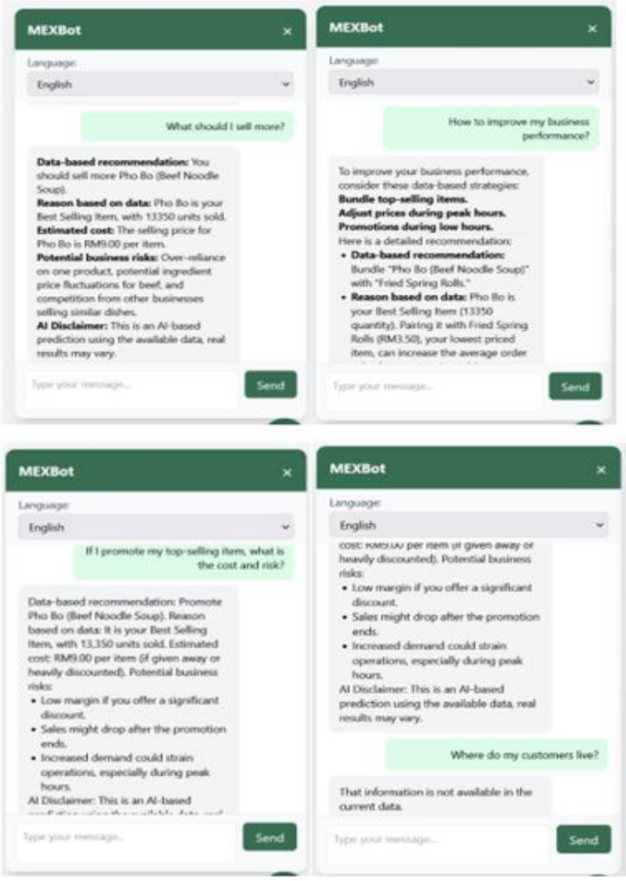


Figure 5. Responses to Innovative Queries

Figure 5 shows that the chatbot not only answers factual questions but is also capable of providing product recommendations and business development ideas based on sales data. The generated responses include bundling potential, cost estimations, and business risk considerations, thereby supporting the research objective of assisting MSMEs in improving business performance through data-driven insights.

In addition, the system includes an AI disclaimer stating that the results are predictive and based on available data, and therefore may differ from real-world conditions.

Testing Results

This section explains the evaluation process of the chatbot system, including functionality testing, intent detection accuracy, similarity score evaluation, assessment using LLM-as-a-Judge, and hallucination testing. All tests were conducted to ensure the accuracy, reliability, and overall quality of the chatbot's responses.

System Functionality Testing

Table 1. System Functionality Testing Results

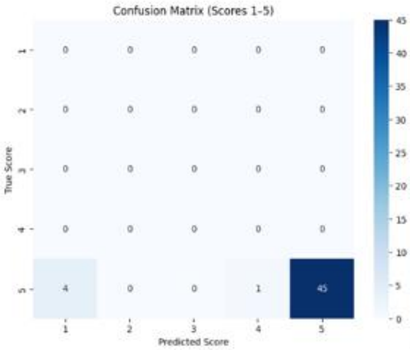
No	Test Case	Input	Expected Output	Actual Output	Status
1	Menu data retrieval	"What is the menu?"	System displays menu list based on merchant_id	System displays correct menu list	Successful
2	Display best-selling product	"What is the best seller?"	System displays item with highest quantity_sold	System displays best-selling item based on CSV data	Successful
3	Display peak ordering hour	"What is the peak hour?"	System displays hour with highest number of orders	Output matches order_hours.csv data	Successful
4	Display monthly sales	"Show monthly sales"	Displays monthly_total_sales, monthly_order_count, and avg_order_value	Output matches monthly_sales.csv	Successful

5	Query not in dataset	"Do I sell spaghetti?"	System informs item is not in the menu	System responds: "No, that item is not in your menu."	Successful
6	Greeting	"Hello"	System responds to greeting	System responds correctly	Successful
7	Unrecognized intent	"Tell me something interesting."	System informs question is not recognized	System provides fallback response	Successful
8	Empty data query	"What is the best seller?" (merchant without data)	System informs data is unavailable	System provides appropriate error message	Successful
9	Gemini API error handling	API error 429	System retries up to 5 times	System resumes after retry	Successful
10	Answer format compliance	Any query	Answer must be concise, clear, and follow prompt rules	Answer follows prompt instructions	Successful
11	Access limitation	"What can affect the peak hour?"	Chatbot responds: "That information is not available in the current data."	System responds that data is unavailable	Successful

The testing results indicate that all core functions—such as menu data retrieval, best-selling product identification, peak ordering hour detection, monthly sales reporting, and performance summaries operate correctly and consistently with the dataset, supported by the performance of the Gemini API.

Intent Detection Accuracy Testing

Intent detection accuracy testing was conducted using 50 sample questions that had been labeled according to their respective intent categories. The evaluation results are presented in a classification report containing precision, recall, F1-score, and support values for each class (scores 1–5), as well as a Confusion Matrix shown in Figure 6.



	precision	recall	f1-score	support
1	0.000	0.000	0.000	0
2	0.000	0.000	0.000	0
3	0.000	0.000	0.000	0
4	0.000	0.000	0.000	0
5	1.000	0.900	0.947	50
accuracy			0.900	50
macro avg	0.200	0.180	0.189	50
weighted avg	1.000	0.900	0.947	50

Figure 6. Confusion Matrix and Intent Evaluation Results

For class 5, the system demonstrates excellent performance, achieving a precision of 1.000, recall of 0.900, and an F1-score of 0.947. This indicates that the system is able to consistently and accurately detect the most frequently used intent category. Overall, the system achieved an accuracy of 90%, meaning that 45 out of 50 questions were correctly classified. The macro average score appears lower because the average calculation includes classes with no data samples, whereas the weighted average is high since the entire weight is dominated by class 5, which shows very strong performance. Thus, the intent detection component can be considered effective for chatbot applications that receive structured question patterns. However, the evaluation could be improved by using a more balanced dataset so that all intent classes can be thoroughly tested.

LLM-as-Judge Evaluation

The primary evaluation was conducted using the LLM-as-a-Judge method with the Gemini 2.0 Flash model. The system was evaluated based on four aspects: Accuracy, Completeness, Conciseness, and Clarity. The use of this method is supported by research indicating that LLM evaluators are capable of assessing other model responses with high consistency and results that closely approximate human judgment. A comparative study shows that LLM-as-Judge can provide stable evaluations with almost no significant differences compared to assessments conducted by human experts (Aftahee et al., 2025). This approach enables evaluation to be carried out in a more objective, faster, and scalable manner compared to manual evaluation using automated scoring scripts with Gemini. The evaluation process follows this structure:

1. **Input:** question, expected answer, actual answer
2. **Output:** scores ranging from 1–5 for Accuracy, Completeness, Conciseness, Clarity, Final Score, and Explanation

Testing was conducted on 10 merchants (out of 100 merchants) due to API token limitations. The final average results can be seen in Table 2.

Table 2. Evaluation Results Using LLM-as-Judge

Metric	Value
Accuracy	4.68
Completeness	4.64
Conciseness	3.80
Clarity	4.68
Final Score	4.66

Based on these results, the system achieved an Accuracy score of 4.68, Completeness of 4.64, Conciseness of 3.80, Clarity of 4.68, and a Final Score of 4.66. The visualization of these evaluation results is presented in Figure 7 in the form of a radar chart to provide a comparative overview of performance across each evaluation metric.

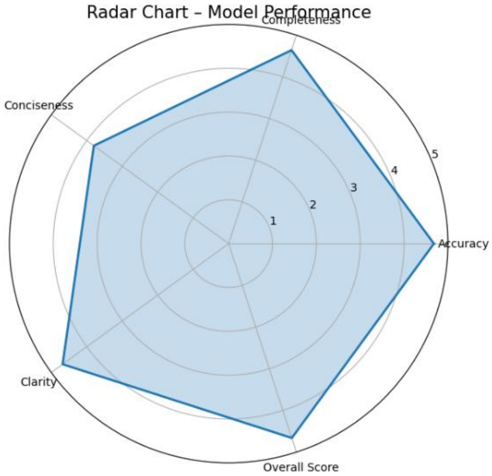


Figure 7. Radar Chart of Evaluation Model Performance

These results indicate that the chatbot demonstrates very high answer quality. The slightly lower Conciseness score is due to the responses tending to be detailed, particularly because of the lengthy CSV data format.

Hallucination Evaluation

The hallucination evaluation was conducted by asking questions about fake menu items, data that does not exist in the CSV file, and ambiguous queries. An example of this testing scenario can be seen in Figure 3.8.

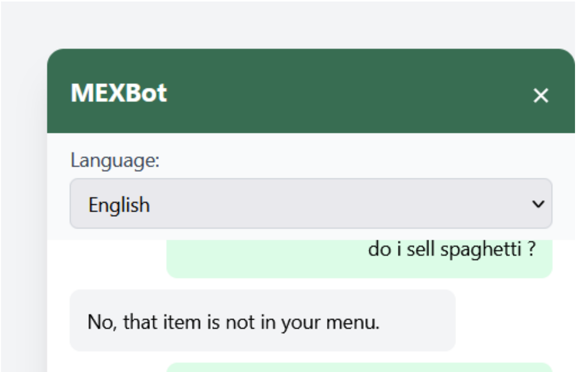


Figure 8. Chatbot Response for Hallucination Testing

When asked:
"Do I sell spaghetti?"

The chatbot responded:
"No, that item is not in your menu."

The results indicate that the system has a low level of hallucination. All factual responses are directly derived from the dataset, and the prompt design restricts the LLM from generating fabricated information.

Discussion

The evaluation results indicate that the chatbot system is capable of providing consistent and relevant responses based on merchant data. Data processing through the ChatbotService module functions effectively, ensuring that the information delivered to users is accurate and aligned with the dataset content. The context-restriction mechanism embedded in the prompt also plays a crucial role in minimizing hallucination. This finding aligns with previous studies suggesting that LLMs become more stable and reliable when provided with structured factual context. In terms of system performance, the chatbot demonstrates strong capability in reading the dataset and generating factual responses. The consistency of its answers reflects that both data processing and prompt engineering have been implemented according to the intended design. The system is able to extract relevant information from the CSV data and present it clearly to users without deviating from the available data source.

Beyond delivering data-based information, the LLM also shows the ability to generate practical insights, including product recommendations and sales strategy suggestions. This capability highlights the system's potential to support micro, small, and medium enterprises (MSMEs), particularly those lacking data analysis expertise, in understanding their business performance and making data-driven decisions. Although the system has been tested using transaction data from merchants within the dataset, it has not yet been evaluated on merchants outside the dataset. Therefore, further validation is necessary to confirm its effectiveness in real-world deployment scenarios. Regarding errors and hallucination, most inaccuracies are influenced by dataset limitations or overly ambiguous user queries. Nevertheless, the overall hallucination rate remains low because responses are strictly constrained to information contained within the CSV file. This controlled generation approach effectively reduces the risk of fabricated or speculative answers.

The evaluation of the chatbot system in this study was conducted entirely using the LLM-as-a-Judge method. This approach was employed to objectively assess response quality across multiple evaluation metrics, including accuracy, completeness, conciseness, and clarity. The results demonstrate stable and consistent scoring, indicating that the LLM-as-a-Judge method is effective for evaluating chatbot performance in response to user queries. Despite its strengths, several limitations remain. The intent detection mechanism is still rule-based, making it less flexible in handling diverse or unexpected user inputs. Additionally, the data used is not real-time, and API quota limitations restricted the scope of testing. These limitations present opportunities for further development, including improving intent classification with machine learning approaches, integrating real-time data processing, and conducting direct user studies with MSME merchants in future research.

CONCLUSION

This study successfully developed a business analysis chatbot system that leverages the capabilities of a Large Language Model to assist MSME owners in understanding their business performance. The system is able to process user questions, identify information needs, and present relevant responses in the form of menu information, sales data, ordering times, and business insights. This approach enables the chatbot to function as a simple analytical assistant that can be accessed by users without technical expertise in data analysis. In addition to providing factual information, the chatbot is also capable of generating recommendations and business development ideas that are easy to understand. This is particularly important because many MSMEs lack data literacy or the resources to conduct independent business analysis. With this capability, the chatbot helps MSMEs identify trends, understand demand patterns, and make more informed decisions. Based on the testing results, the system demonstrates stable performance and consistent responses aligned with the context of user queries. The logic-restriction mechanism implemented through prompt design effectively reduces misinformation, ensuring that the chatbot provides answers that remain consistent with the available data. Overall, this study has achieved its objective of delivering a business chatbot system that is informative, user-friendly, and beneficial for business decision-making.

REFERENCES

Aftahee, S., Farhad, A. F. M., Mallik, A., Dhar, R., Karim, J., Noor, N. B., & Solaiman, I. A. (2025). *Assessing the Reliability of Large Language Models in the Bengali Legal Context: A Comparative Evaluation Using LLM-as-Judge and Legal Experts* (arXiv:2511.05627). arXiv. https://doi.org/10.48550/arXiv.2511.05627

Akhtar, Z. B. (2024). Unveiling the evolution of generative AI (GAI): A comprehensive and investigative analysis toward LLM models (2021–2024) and beyond. *Journal of Electrical Systems and Information Technology*, *11*(1), 22. https://doi.org/10.1186/s43067-024-00145-1

Alsadoun, A. A., & Alnasser, A. N. (2025). Role of artificial intelligence chatbot marketing in enhancing customer satisfaction and loyalty in digital shopping experiences. *Int. J. Innov. Res. Sci. Stud*, *8*(1), 1902–1909.

Anggraini, J., Oktadinna, N. K., & Martini, M. (2025). Transformasi sumber daya manusia dalam era industri 5.0: Tantangan dan peluang pengembangan karyawan. *Journal of Business, Finance, and Economics (JBFE)*, *6*(1), 252–266.

Ardiansyah, R. H., & Sulaksono, A. G. (2023). Layanan pelanggan berbasis Natural Language Processing melalui chatbot pada aplikasi pesan. *Journal of Information System and Application Development*, *1*(1), 29–37.

Chitturu, S., Lin, D.-Y., Sneader, K., Tonby, O., & Woetzel, J. (2017). Artificial intelligence and Southeast Asia's future. *Singapore Summit*, 1–40.

Fahmi, S. (2024). Pemanfaatan teknologi AI untuk menunjang pemasaran produk UMKM di Kota Malang. *BERDAYA EKONOMI: Jurnal Pengabdian Masyarakat*, *2*(2), 91–103.

Festiyed, -, Tanjung, Y. I., & Fadillah, M. A. (2024). ChatGPT in Science Education: A Visualization Analysis of Trends and Future Directions. *JOIV : International Journal on Informatics Visualization*, *8*(3–2), 1614–1624. https://doi.org/10.62527/joiv.8.3-2.2987

Fitriani, A., Sulaeman, D., Firmansyah, R., Ekonomi, F., & Stembi, S. (2022). Efektivitas chatbot sebagai media komunikasi bisnis antara penjual dan pembeli pada Marketplace. *Jurnal Disrupsi Bisnis*, *5*(3), 189–196.

Godwin, G., Junaedi, S. R. P., Hardini, M., & Purnama, S. (2024). Inovasi Bisnis Digital untuk Mendorong Pertumbuhan UMKM melalui Teknologi dan Adaptasi Digital. *ADI Bisnis Digital Interdisiplin Jurnal*, *5*(2), 41–47. https://doi.org/10.34306/abdi.v5i2.1172

Hadi, M. U., Tashi, Q. A., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). *Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects*. https://doi.org/10.36227/techrxiv.23589741.v3

Harahap, A. R., Munthe, C., Hariri, M. F., & Lubis, S. P. (2025). Peran AI dalam UMKM: Bagaimana Kecerdasan Buatan Membantu UMKM Bertahan dan Berkembang di Era Digital. *IKRAITH-EKONOMIKA*, *8*(1), 409–419.

Harisi, M. R., & Hiwono, E. M. (2024a). Pengaruh Chatbot Usage terhadap Customer Satisfaction. *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi*, *1*(2), 66–73.

Harisi, M. R., & Hiwono, E. M. (2024b). Pengaruh Chatbot Usage terhadap Customer Satisfaction. *Jurnal Ilmiah Manajemen Ekonomi Dan Akuntansi*, *1*(2), 66–73.

Jurafsky, D., & Martin, J. H. (2000). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. https://aclanthology.org/anthology-files/pdf/J/J00/J00-4006.pdf

Katti, S. W. B., & Mutmainah, M. (2020). Penyuluhan Pentingnya Pembukuan Umkm Sederhana Dan Tahapan Penyusunan Studi Kelayakan Usaha. *Jurnal Daya-Mas*, *5*(2), 58–60.

Kee, D. M. H., Mohd Rusdi, S. A., Mokhtar, S. S., Muhamad Ridzuan, S. A., & Amni Abdullah, S. A. (2021). How Grab Becomes One of the Greatest Start-ups in South-East Asia. *International Journal of Tourism and Hospitality in Asia Pasific*, *4*(1), 73–81. https://doi.org/10.32535/ijthap.v4i1.1023

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2022). Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*, *54*(10s), 1–41.

Latifah, D. Z., Safitri, K. Y., Aisyah, N., Jovena, S., Arsamitha, W. G. M., Syakila, Z. A., & Badri, M. (2026). Transformasi Digital UMKM melalui Inovasi Platform Bisnis Berkelanjutan. *Jurnal Penelitian Ilmiah Multidisipliner*, *2*(04), 977–986.

Maria, S. (2024). *Compass: Large Multilingual Language Model for South-east Asia* (arXiv:2404.09220). arXiv. https://doi.org/10.48550/arXiv.2404.09220

Mijwil, M. M., Aggarwal, K., Sonia, S., Al-Mistarehi, A. H., Alomari, S., Gök, M., Zein Alaabdin, A. M., & Abdulrhman, S. H. (2022). Has the Future Started? The Current Growth of ArtificialIntelligence, Machine Learning, and Deep Learning. *Iraqi Journal for Computer Science and Mathematics*, *3*(1), 13.

Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2025). *Large Language Models: A Survey* (arXiv:2402.06196). arXiv. https://doi.org/10.48550/arXiv.2402.06196

Novitasari, A. T. (2022). Kontribusi UMKM Terhadap Pertumbuhan Ekonomi Era Digitalisasi Melalui Peran Pemerintah. *JABE (Journal of Applied Business and Economic)*, *9*(2), 184–204. https://doi.org/10.30998/jabe.v9i2.13703

Pichai, S., & Hassabis, D. (2023, December 6). *Introducing Gemini: Our largest and most capable AI model*. Google. https://blog.google/innovation-and-ai/technology/ai/google-gemini-ai/

Rane, N., Choudhary, S., & Rane, J. (2024). Gemini versus ChatGPT: Applications, performance, architecture, capabilities, and implementation. *Journal of Applied Artificial Intelligence*, *5*(1), 69–93. https://doi.org/10.48185/jaai.v5i1.1052

Russell, S., Norvig, P., & Intelligence, A. (1995). A modern approach. *Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs*, *25*(27), 79–80.

Suryanto, T. L. M., Wibawa, A. P., Hariyono, H., & Nafalski, A. (2023). Evolving conversations: A review of chatbots and implications in natural language processing for cultural heritage ecosystems. *International Journal of Robotics and Control Systems*, *3*(4), 955–1006.

Sutisna, Hidayat, S., Affandi, N., Komaruddin, M., & Pratama, A. G. (2025). Digital Technology Literacy and Digital Economy Literacy as an Effort to Improve Business Sustainability in Banten's Micro and Small Industries in the Era of Business Uncertainty. *Journal of Computer Science*, *21*(4), 761–778. https://doi.org/10.3844/jcssp.2025.761.778

Sutisna, & Raudhan, M. N. (2025). Strategi Pengambilan Keputusan Berbasis Data Analytics pada UMKM untuk Meningkatkan Penjualan. *Data Sciences Indonesia (DSI)*, *5*(2), 38–49. https://doi.org/10.47709/dsi.v5i2.6768

Utami, R. S., & Sasmita, M. A. (2022). Strategi umkm dalam menghadapi pasar di masa pandemi Covid-19 (umkm dalam bidang kuliner di wilayah Kecamatan Tanjung Morawa). *Jurnal Widya*, *3*(1), 81–94.

World Bank. (2021). *World Bank SME Finance* [Text/HTML]. World Bank. https://www.worldbank.org/en/topic/smefinance