JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

INTELLIGENT CHATBOT FOR CUSTOMER SUPPORT AND SERVICE

[1]Tejas Shigwan, [2]Saurabh Kadam, [3]Harsh Ghosalkar, [4]Arti Devmane

[1]Student, [2]Student, [3]Student, [4]Faculty of Information Technology
[1]Department of Information Techonology,
[1]Shivajirao S. Jondhale College of Engineering, Thane, India

Abstract : This project presents a web application leveraging Streamlit and the Gemini API to offer two key functionalities: a Vision Application and a Q&A Chat Application. The Vision Application allows users to upload images and input prompts, generating descriptive responses based on the content of the image and prompt. The Q&A Chat Application enables users to ask questions, with the Gemini API providing responses. Both applications offer real-time interaction and display chat history, showcasing the integration of AI technologies for image analysis and natural language processing in a user-friendly web environment

I. Introduction

Our project presents a user-friendly web application that leverages AI technology to provide two key functionalities: a Vision Application and a Q&A Chat Application. Powered by the Gemini API, the Vision Application allows users to upload images and input prompts, generating descriptive responses based on image content. The Q&A Chat Application enables real-time interaction, where users can ask questions and receive responses generated by the Gemini API. Both applications offer seamless integration, providing an intuitive interface for users to engage with advanced AI capabilities.

II. LITERATURE SURVEY -

2.1 OBJECTIVE –

1. Develop Intelligent Applications: The primary objective of this project is to develop two intelligent applications: the Vision Application and the Q&A Chat Application, leveraging cutting-edge AI technologies to provide advanced functionalities in image understanding and question-answering.

2. enhance user experience: the project aims to enhance user experience by providing intuitive and user-friendly interfaces for interacting with the ai-powered applications. this includes features such as easy input prompts, image upload capabilities, and real-time responses.

3. utilize ai models: utilize state-of-the-art ai models, including the gemini image model and the gemini large language model (llm), to power the vision and q&a chat applications, respectively. these models are integrated using the google generativeai api to ensure high-quality responses and accurate predictions.

4. provide image understanding: enable the vision application to analyze and interpret images based on natural language prompts, providing descriptive responses about the content of the uploaded images. this includes identifying objects, scenes, and other relevant information within the images.

5. enable question-answering: enable the q&a chat application to engage in conversational interactions with users, answering queries and providing relevant information in real-time. the chatbot should be capable of understanding natural language queries and generating accurate responses.

6. ensure efficiency and reliability: ensure that both applications provide prompt, accurate, and reliable responses to user queries, minimizing wait times and maintaining consistency in the quality of interactions. this includes optimizing the performance of ai models and handling user inputs effectively.

2.2 Literature Review -

1. Customer Support Chatbot - This project focuses on developing a chatbot for customer support applications, similar to our Q&A Chat Application. Drawback - Lack of Image Understanding: Unlike our Vision Application, Project A's chatbot does not incorporate image understanding capabilities, limiting its ability to assist users with image-related queries.

2. Image Recognition Chatbot - Image recognition and understanding, akin to our Vision Application. Drawbacks - Lack of Conversational Abilities: While Project B excels in image recognition, it may lack conversational abilities and the capacity to engage users in dynamic interactions, as seen in our Q&A Chat Application.

3. Hybrid Chatbot system - This project combines elements of both image recognition and conversational AI, resembling a hybrid of our Vision and Q&A Chat Applications. Drawbacks - Complexity and Overhead: Developing and maintaining a hybrid system like Project C may introduce complexity and overhead in terms of integration, model management, and user experience design.

Traditional vision-based systems - Traditional vision-based systems rely on handcrafted features and rule-based algorithms for image understanding. Drawback - These systems often struggle with complex images and require extensive manual feature engineering, making them less adaptable to diverse datasets.

III. METHODS AND MATERIALS :

3.1 Methodology -

1. Problem Formulation:
- Define the problem statement and objectives of the project, focusing on developing two distinct applications: the Vision Application and the Q&A Chat Application. Clearly outline the tasks each application is expected to perform.

2. Data Collection:
- Identify and collect relevant datasets for training and testing the AI models used in both applications. This includes image datasets for the Vision Application and conversational datasets for the Q&A Chat Application.

3. Model Selection:
- Research and select appropriate AI models and algorithms for image understanding and natural language processing tasks. Consider factors such as model performance, computational efficiency, and compatibility with the project requirements.

4. Model Training:
- Train the selected AI models using the collected datasets. Fine-tune the models as necessary to optimize performance and adapt them to the specific tasks of image understanding and question answering.

5. Application Development:
- Implement the Vision Application and Q&A Chat Application using Streamlit for the user interface. Integrate the trained AI models into the applications to enable real-time inference and interaction with users.

6. User Interface Design:
- Design intuitive and user-friendly interfaces for both applications, focusing on simplicity, clarity, and ease of use. Incorporate features such as text input fields, image uploaders, and interactive buttons to facilitate user interaction.

7. Testing and Evaluation:
- Conduct comprehensive testing of the Vision Application and Q&A Chat Application to assess their functionality, performance, and accuracy. Evaluate the applications using a variety of test cases and scenarios to ensure robustness and reliability.

8. Performance Metrics:
- Define evaluation metrics to measure the performance of the applications, such as accuracy, precision, recall, and F1-score. Use these metrics to quantify the effectiveness of the AI models and the overall performance of the applications.

9. User Feedback and Iteration:
- Gather feedback from users and stakeholders through user testing and surveys. Use this feedback to identify areas for improvement and iterate on the design and functionality of the applications accordingly.

10. Deployment and Documentation:
- Deploy the Vision Application and Q&A Chat Application for real-world use, ensuring that they are accessible and functional. Document the development process, including code, configurations, and dependencies, for future reference and maintenance.

11. Maintenance and Support:
- Provide ongoing maintenance and support for the deployed applications, addressing any issues or bugs that arise and incorporating updates and enhancements as needed to improve performance and user experience.

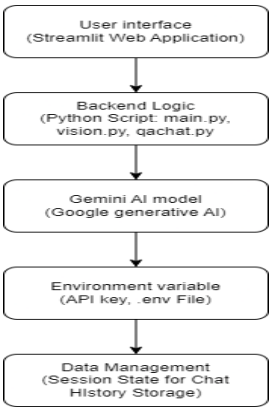

Figure 1: Block Diagram

3.2 Materials -

Hardware specification:
1. Processor: Intel i3 or above
2. RAM: 4GB or above
3. HDD: 500GB or above

Software specification:
1.Operating System: Windows 10 or above
2.IDE - Visual studio
3. Language: python

IV. RESULTS AND DISCUSSION

4.1 Results

Vision Application
- Gemini Image Demo:
  - Users can input a prompt and upload an image.
  - Upon submission, the system generates responses based on the input prompt and the content of the uploaded image.
  - The response is displayed to the user.

Q&A Chat Application
- Gemini LLM Application:
  - Users can input a question.
  - Upon submission, the system generates responses based on the input question.
  - The conversation history is displayed, including both user inputs and bot responses.

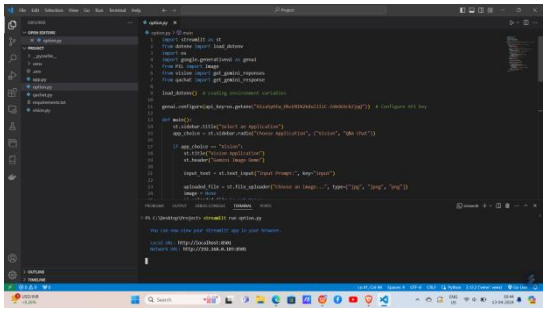

Figure 2. Start the project

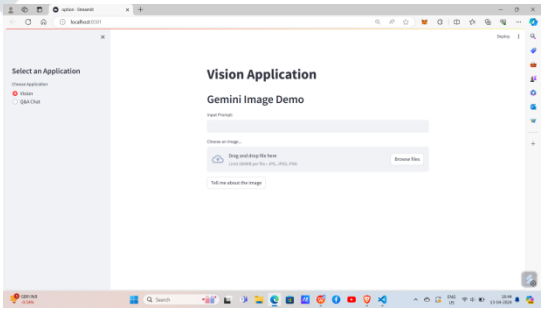

Figure 3. Vision application interface

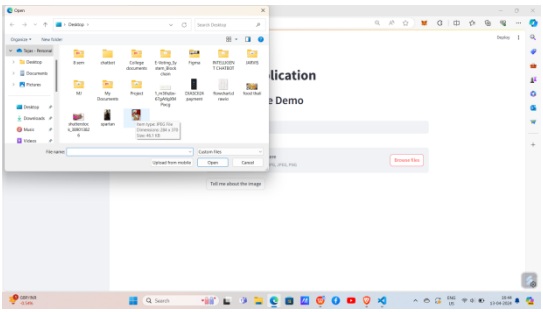

Figure 4. Image improting

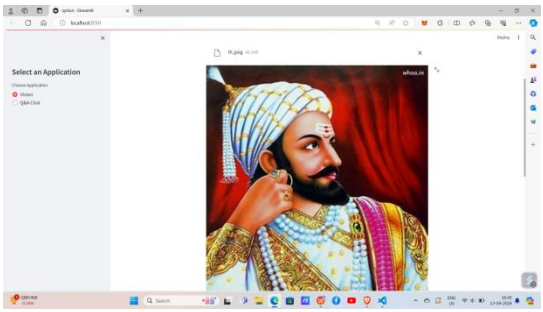

Figure 5. Image import successfully

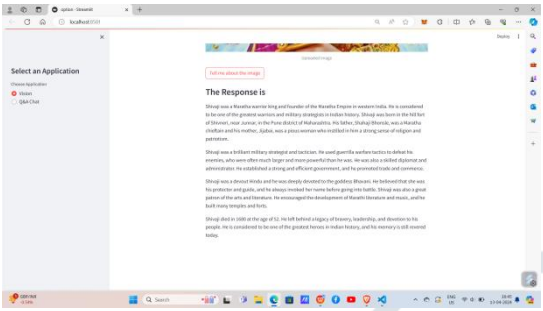

Figure 6. Image description design

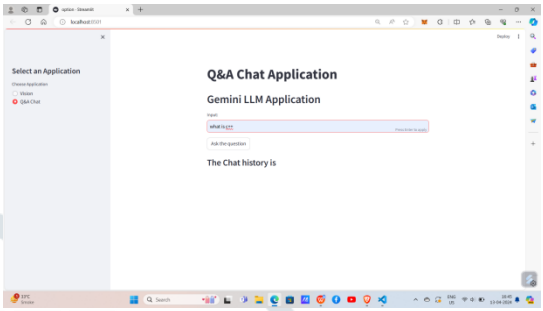

Figure 7. Q&A Chat application interface

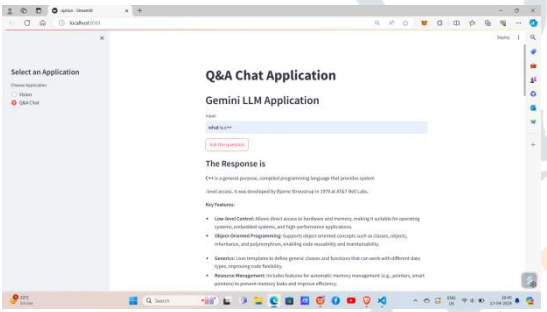

Figure 8. Quention & answer part

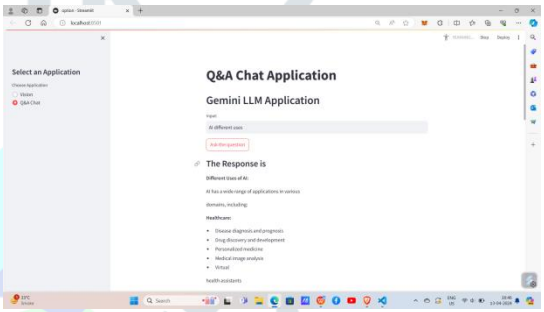

Figure 9. Answer generation

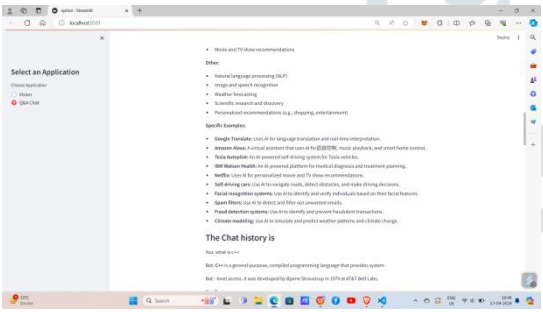

Figure 10. Chat history

4.2 Discussion

4.2.1 Vision Application -

Effectiveness of Gemini Image Demo:

The Gemini model successfully generates responses based on both textual prompts and visual content. This demonstrates the capability of the model to understand and interpret images in addition to text.The system's ability to provide meaningful responses depends on the quality and relevance of the input prompt and the uploaded image. Users may need guidance on how to formulate effective prompts to receive accurate responses.

4.2.2 Q&A Chat Application -

Performance of Gemini LLM Application:

The Gemini model provides responses to user questions, demonstrating its ability to understand and generate human-like text.
The system's responses vary in accuracy and relevance depending on the complexity and specificity of the questions asked. Simple and direct questions tend to yield more accurate responses compared to ambiguous or multi-faceted questions.
Continuous improvement in the training data and fine-tuning of the model could enhance the accuracy and relevance of responses over time.

4.2.3 Overall System Performance

User Experience:
The user interface provided by Streamlit offers a user-friendly experience, allowing users to interact with the applications easily. However, there may be limitations in terms of scalability and responsiveness, particularly as the user base and data volume increase.

V. ACKNOWLEDGMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose constant guidance, supervise and encouragement crowned our efforts with success. We consider our privilege to express gratitude and respect to all who guided us in the completion of the major project. We convey our gratitude is our project guide *Prof. Arti Devmane* for her timely help, support, encouragement, constructive suggestion, information and many innovative ideas in carrying out the project. We would also like to extend our gratitude to our respected Principal *Dr. P. R. Rodge*, and our Head of Department *Dr. Savita Sangam*, for their support and encouragement throughout the development of this major project.

REFERENCES

[1] Brady D. Lund & Ting wang, February 2023, "Chatting about chatGPT" is Article in Library Hi Tech News.

[2] Ivan Martins De Andrade and Cleonir Tumelero Business School, Positivo University, Curitiba, Brazil, 7 February 2022 "Increasing customer service efficiency through artificial intelligence chatbot".

[3] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020) "Language models are few-shot learners." Advances in Neural Information Processing Systems, , 33, 1877-1901.

[4] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017) "Attention is all you need." Advances in Neural Information Processing Systems, 30, 5998-6008.

[5] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019), Language models are unsupervised multitask learners. OpenAI Blog, 1(8), 9.

[6] Devlin, J., Chang, M. W., Lee, K., & Toutanova, (2018) "K. BERT: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805 .

[7] Rennie, J., Marcheret, E., Mroueh, Y., Ross, J., & Goel, V. (2017) Self-critical sequence training for image captioning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition., 7008-7024.

[8] Brown, T. B., Bagnell, D., & Sutskever, I., (2020) Adaptive Attention Span in Transformers. arXiv preprint arXiv:1905.07799.