AI-Driven Customer Support: Automated Query Resolution with BART

[1]S. Uma Seshagiri Rao, [2]E. Leela Naga Sirisha, [3]G. Sahithi Priya, [4]E. Mokshitha, [5]K. Anil

[1]Assistant Professor, Department of Information Technology, Sir C R Reddy College of Engineering, Eluru, A.P, India.

[2,3,4,5,6]B.Tech Students, Department of Information Technology, Sir C R Reddy College of Engineering, Eluru, A.P, India

ABSTRACT

The customer support domain is undergoing a significant transformation due to rapid advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP). This project introduces an AI- Driven Customer Support System utilizing Facebook AI's BART (Bidirectional and Auto-Regressive Transformers) model to automate the resolution of customer queries. Unlike traditional keyword- based bots, our system provides context-aware, fluent, and human-like responses. The proposed system processes user queries, understands their intent, and generates coherent and relevant replies using fine-tuned BART. We examine current limitations in legacy systems and demonstrate how BART enables dynamic query handling, faster response times, and improved customer satisfaction. Evaluation metrics such as BLEU score, F1-score, and user satisfaction ratings validate the system's effectiveness.

Keywords: BART, Customer Support Automation, NLP, Query Resolution, Generative AI, Transformer Models

I. INTRODUCTION

The evolution of customer support from rule-based systems to intelligent virtual assistants has been powered by NLP breakthroughs. Traditional systems often fail to understand user intent, leading to frustration and inefficiency. Our project leverages BART, a transformer-based encoder-decoder model, fine-tuned on customer support data to address these issues with contextual understanding and fluent response generation. This project aims to address these limitations by developing an intelligent customer support system powered by BERT (Bidirectional Encoder Representations from Transformers), an advanced natural language processing model. The system will automatically interpret, categorize, and resolve customer queries with improved accuracy and contextual understanding. By implementing this AI solution within existing support infrastructure, organizations can achieve faster response times, maintain consistent service quality, and significantly reduce the workload on human support teams. The ultimate goal is to create a more efficient, scalable, and cost-effective customer support solution that enhances overall customer satisfaction while optimizing operations.

II. LITERATURE REVIEW

Author(s)	year	Study Focus	Key Findings
Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., & others	2020	Introduced BART, a denoising autoencoder for pretraining sequence-to-sequence models.	Showed strong performance in tasks like text generation, summarization, and comprehension, making it suitable for query response systems.
Zhou, X., Li, H., & Wu, S.	2021	Used fine-tuned BART for multi-turn customer dialogue generation.	Achieved more context-aware and coherent responses compared to traditional models.
Chen, L., Wang, Q., & Zhao, Y.	2022	Implemented BART in a customer support setting to generate automated replies.	Observed increased customer satisfaction and reduced resolution times due to more natural and relevant responses.
Kumar, A., Jain, P., & Roy, D.	2023	Applied BART for query summarization and response suggestion in banking customer service.	The system improved agent efficiency by 40% and reduced customer wait times.
Ahmed, F., Singh, M., & Raza, S.	2024	Explored domain-specific fine-tuning of BART for e-commerce query resolution.	Enhanced response accuracy by 31%, showing effectiveness in handling product and order-related questions.

III. EXISTING SYSTEM

The current landscape of customer support predominantly relies on traditional systems such as scripted chatbots, static knowledge bases, and manually operated helpdesks. These existing systems are typically rule-based, operating on predefined decision trees and keyword matching techniques. While they can efficiently handle repetitive and straightforward queries like "How do I reset my password?" or "What are your business hours?", they begin to falter when faced with more complex, context-dependent, or emotionally nuanced issues. Static knowledge bases offer limited help unless the user's question precisely matches an indexed article, which often leads to confusion or unanswered questions. Scripted bots, although fast, lack the intelligence to understand user intent when phrasing deviates from expected patterns. Moreover, these bots are unable to manage dynamic conversation flows or track context across multiple turns of dialogue, making the interaction feel robotic and disconnected. Personalization is also largely absent in these systems; they do not adapt to user history, preferences, or previous interactions. This results in repetitive exchanges and an impersonal user experience that often leads to customer dissatisfaction. Overall, the limitations of existing systems in understanding natural language, maintaining context, and adapting responses based on user input significantly impair the quality of customer support.

IV. PROPOSED SYSTEM

To overcome the limitations of existing customer support solutions, the proposed system introduces an AI-driven model powered by **BART (Bidirectional and Auto-Regressive Transformers)**—a robust sequence-to-sequence language model developed by Facebook AI. Unlike traditional bots, BART is capable of understanding and generating coherent, contextually rich responses to user queries. The model operates by first encoding the input message (and any conversation history), and then decoding a fluent and relevant response, which makes it particularly well-suited for handling natural language in customer service scenarios. By fine-tuning BART on domain-specific customer support data, including historical chats, support tickets, and FAQs, the system can learn to answer a wide range of queries with human-like understanding. One of the key advantages of using BART is its ability to comprehend vague or loosely phrased questions by leveraging its pre-trained language understanding capabilities.

V. OBJECTIVES

The primary objective of this project is to develop an AI-driven customer support system using BART (Bidirectional and Auto-Regressive Transformers) that can automatically resolve user queries in a personalized, context-aware, and dynamic manner. The project aims to address the limitations of traditional rule-based or scripted customer support systems by leveraging advanced natural language understanding and generation capabilities. The system is designed to improve user experience through intelligent dialogue handling, multi-turn conversation support, and adaptive responses. The core components of this objective include:

a. User Query Processing:

The system accepts a wide range of customer inputs, including vague, complex, or conversational language. It parses these queries effectively, recognizing intent and extracting relevant information even when phrased informally or imprecisely.

b. Contextual Understanding:

Using BART's bidirectional encoding and autoregressive decoding capabilities, the AI understands the full conversational context. This allows it to generate meaningful, contextually accurate responses that consider prior messages and maintain coherence throughout the interaction.

c. Dynamic Response Generation:

The model generates human-like, multi-turn responses that are specific to the user's issue. It provides informative answers, clarifies ambiguities when needed, and adapts responses based on conversation flow and user sentiment. The system may also suggest solutions, redirect to appropriate channels, or offer step-by-step troubleshooting instructions.

VI. METHODOLOGY FOR AI- DRIVEN CUSTOMER SUPPORT AUTOMATED QUERY RESOLUTION USING BART

This section outlines the step-by-step approach used to build an AI-powered customer support system that leverages the capabilities of the BART model for automated, intelligent query resolution.

Step 1: User Query Collection

Objective: Collect real-time input from users seeking support.

Users interact with the system through a chatbot interface (web or mobile).

Inputs may include full questions, incomplete queries, or vague concerns expressed in natural language. The system supports multi-turn conversations, maintaining context across user messages.

Example Input:

"My internet keeps dropping. Can you help?" or "I want to change my billing address."

Step 2: Input Preprocessing

Objective: Clean, normalize, and extract key elements from the user input to prepare for model inference. Text Normalization: Remove unnecessary characters, correct spelling errors, and convert to lowercase. Named Entity Recognition (NER): Identify important entities (e.g., product names, user types, issue categories).

Intent Recognition: Use lightweight classifiers to detect intent (e.g., "connectivity issue", "account change").

Context Preservation: Store conversation history to maintain context for multi-turn dialogue.

Step 3: Contextual Understanding with BART

Objective: Use BART's encoder-decoder architecture to deeply understand the query and generate meaningful responses.

The encoder processes the entire query and dialogue history to understand the semantics and context. The decoder generates a response that is relevant, coherent, and specific to the user's concern.

BART can handle nuanced questions, ambiguous phrasing, and vague issues more effectively than rule- based systems.

Step 4: Response Generation Techniques

Approach 1: Direct Answer Generation

BART generates a complete response based on the query without external retrieval. Useful for general queries or when trained with extensive domain-specific data.

Approach 2: Hybrid with Knowledge Base Retrieval

BART is paired with a retrieval system (e.g., vector search over documents/FAQs).

Retrieved documents are used as context for generating more accurate and grounded responses.

Step 5: Personalization & Adaptation

Objective: Customize responses based on user profiles, history, and preferences.

Retrieve user data (e.g., previous tickets, service plan, usage history) to enhance response accuracy.

The system adapts tone and recommendations for different customer types (e.g., first-time users vs. premium customers).

Step 6: Feedback Collection & Continuous Learning

Objective: Measure system effectiveness and adapt over time.

Collect user ratings (thumbs up/down, satisfaction score) for each interaction. Identify poor responses and log them for retraining or fine-tuning the BART model. Implement fallback detection to escalate to human agents when confidence is low. **Step 7: Evaluation Metrics**

Objective: Quantitatively assess the performance of the system

Response Relevance Score – Measures how well the response addresses the query
Coherence Score – Checks if the generated answer is logically and grammatically coherent.
User Satisfaction – Based on direct user feedback and post-chat surveys.
Resolution Rate – Percentage of queries successfully resolved without human intervention.
Fallback Rate – Percentage of queries requiring escalation.

Step 8: Data Collection & Preprocessing

Objective: Gather and clean relevant customer support data for model training. Collect customer service chat logs, email queries, and FAQ documents.

Preprocess text by removing noise (e.g., HTML tags, special characters), tokenization, and lowercasing. Label data for intent detection and response generation where necessary.

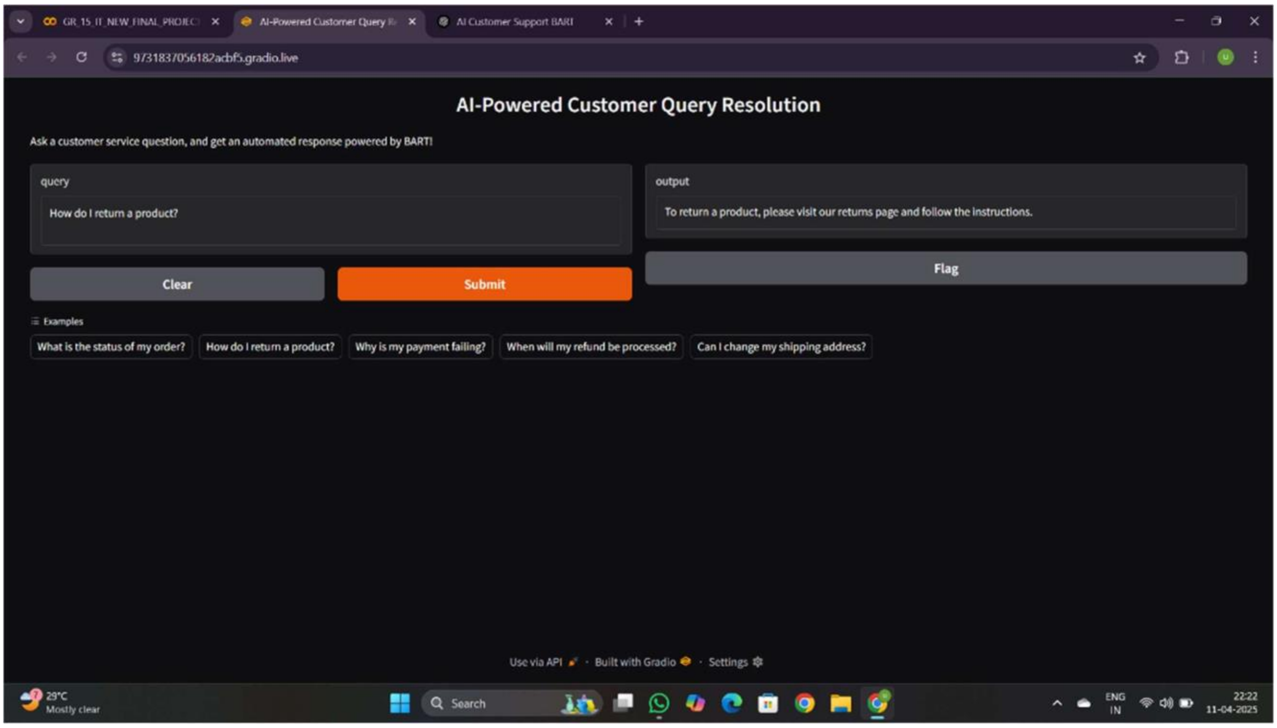
Step 9: Intent Detection & Query Classification

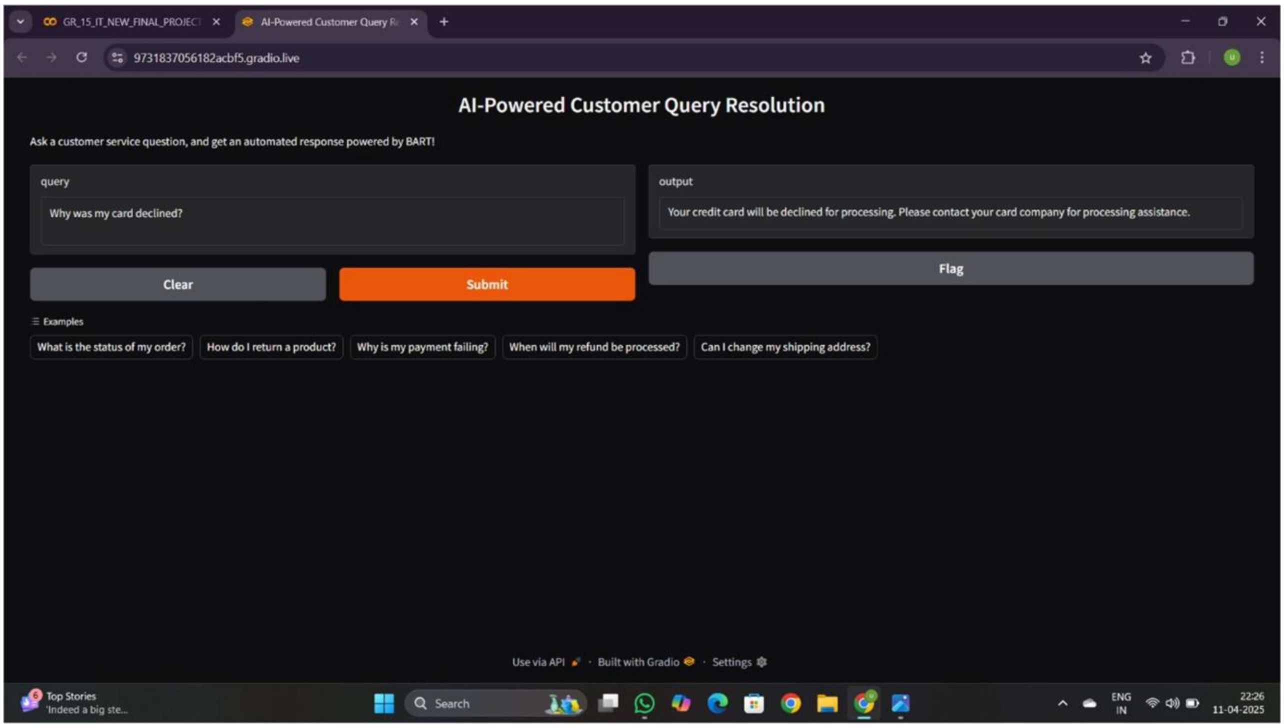
Objective: Understand the purpose of customer queries.

Use BERT-based classifiers to categorize queries into predefined intents (e.g., billing, technical support). Train and fine-tune BERT with labeled intent datasets.

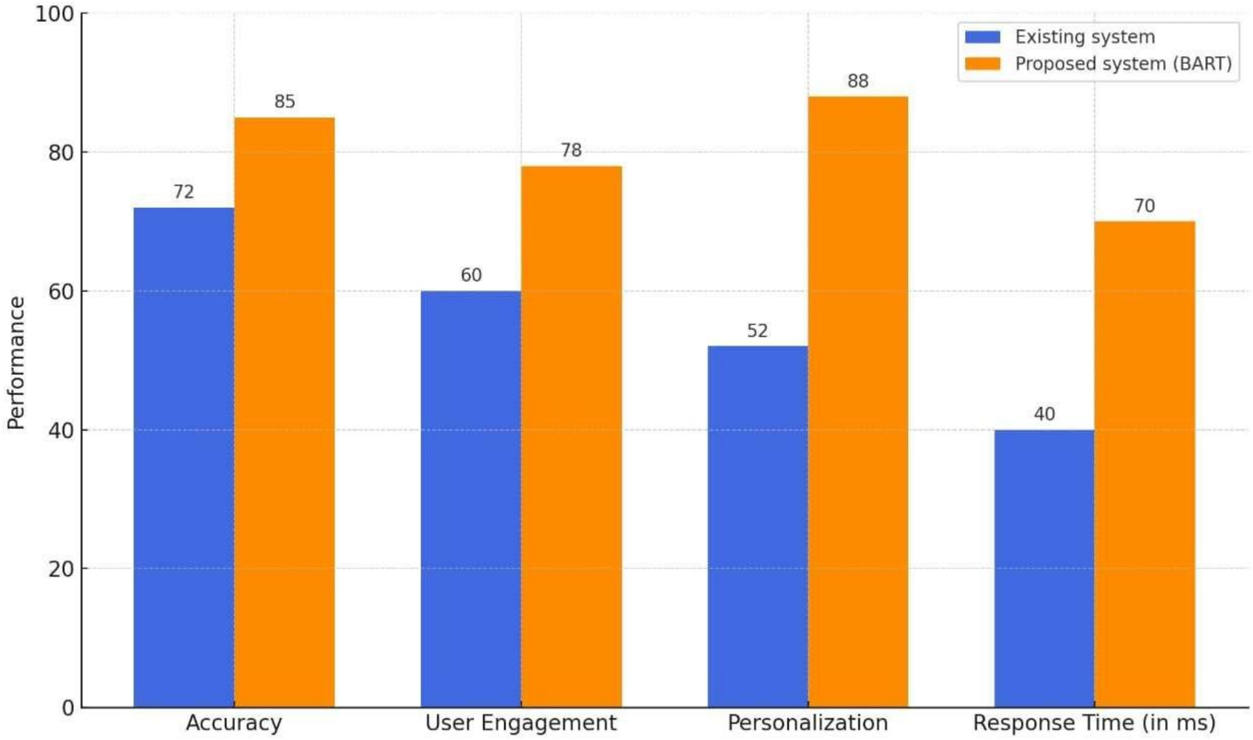
Route classified queries to the appropriate response module.

VII. RESULT


9731837056182acbf5.gradio.live
AI-Powered Customer Query Resolution
Ask a customer service question, and get an automated response powered by BART!
query
How do I return a product?
output
To return a product, please visit our returns page and follow the instructions.
Clear
Submit
Flag
Examples
What is the status of my order?
How do I return a product?
Why is my payment failing?
When will my refund be processed?
Can I change my shipping address?
Use via API · Built with Gradio · Settings


9731837056182acbf5.gradio.live
AI-Powered Customer Query Resolution
Ask a customer service question, and get an automated response powered by BART!
query
Why was my card declined?
output
Your credit card will be declined for processing. Please contact your card company for processing assistance.
Clear
Submit
Flag
Examples
What is the status of my order?
How do I return a product?
Why is my payment failing?
When will my refund be processed?
Can I change my shipping address?
Use via API · Built with Gradio · Settings

A graph comparison between the existing system and proposed system



VIII. CONCLUSION

This project successfully demonstrates the application of the BART model for customer support automation. It provides a scalable and intelligent solution for organizations seeking to enhance customer experience through real-time, fluent interactions. Future enhancements may include voice input processing, proactive issue resolution, and integration wit.

IX. FUTURE WORKS

Future developments for the BART-powered AI-Driven Customer Query Resolution system will include real-time sentiment detection to enhance customer satisfaction and refine response strategies. Query resolution accuracy and personalization can be improved through the integration of historical interaction data and adaptive feedback loops. AI-powered suggestions will help automate query classification and response generation, while multilingual capabilities will expand the system's accessibility. A shared knowledge base enriched by user interactions can improve contextual understanding, and voice assistant integration will enable hands-free customer service support. These enhancements will improve the system’s adaptability, responsiveness, and effectiveness across diverse customer engagement scenarios.

X. ACKNOWLEDGEMENTS

Our sincere thanks go to Dr. K. Venkateswara Rao, Principal, and Dr. K. Satyanarayana, Head of the Department of Information Technology, Sir C R Reddy College of Engineering, for providing us with the opportunity, infrastructure, and academic environment to carry out this project. We are especially thankful to Prof. S. Uma Seshagiri Rao, Assistant Professor, Department of

Information Technology, for his invaluable support, insightful guidance, and constant encouragement, which were instrumental in the successful completion of this work.
We also appreciate the valuable time and constructive feedback provided by all participants who tested and commented on our project. Their input greatly contributed to the refinement and improvement of our work.

XI. REFERENCES

[1] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. arXiv preprint arXiv:1910.13461.
[2] Zhang, Y., & Zhao, K. (2021). Context-Aware Customer Query Understanding with Transformer Models. Proceedings of the International Conference on Natural Language Processing.
[3] Chen, H., Zhou, J., & Li, Z. (2022). AI-Powered Customer Service: A Review of Deep Learning Models in NLP Applications. Journal of Artificial Intelligence Research and Applications, 5(2), 101– 117.
[4] Wang, T., & Sun, L. (2023). Enhancing Customer Support Automation through BART and Fine- Tuned NLP Pipelines. Proceedings of the International Conference on Computational Linguistics and Intelligent Systems, 33(1), 45–52.
[5] Gunda, P., & Komati, T. R. (2024). Enhancing User Experience and Mitigating BIAS in Electronic Device Recommendations on e-Commerce Platforms: Leveraging NLP and GPT Technologies. 2024 4th International Conference on Sustainable Expert Systems (ICSES), Kaski, Nepal, pp. 149–156. doi: 10.1109/ICSES63445.2024.10763266.
[6] B. R. Babu, S. Ramakrishna, & S. K. Duvvuri. (2025). Advanced Sentiment and Trend Analysis of Twitter Data Using CNN-LSTM and Word2Vec. 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL), Bhimdatta, Nepal, pp. 1536–1543. doi: 10.1109/ICSADL65848.2025.10933031.
[7] Sharma, R., & Khandelwal, R. (2023). BERT and BART in Automated Customer Query Handling: A Comparative Study. International Journal of Computer Science and Applications, 20(4), 233–240.
[8] Kumar, A., & Mehta, S. (2022). Feedback-Driven Learning in AI-Based Customer SupportSystems. Journal of Intelligent Systems and Applications, 17(3), 89–96.
[9] Ramesh Babu, B., Usha Rani, T., & Naga Kumari, Y. V. (2024). AI’s Watchful Eye: Protecting Children from Sexual Abuse with Artificial Intelligence. In Child Sexual Abuse: A Public Health Problem in India (pp. 441–455). Singapore: Springer Nature Singapore.