

NLP-Based Intelligent Chatbots

Vishnu Babalsure¹ and J. R. Mahajan²

U.G. Student, Department of Artificial Intelligence & Data Science¹

HOD, Department of Artificial Intelligence & Data Science²

Adsul's Technical Campus, Chas, Ahillyanagar, India

Abstract: Chatbots have become an important application of Artificial Intelligence (AI), enabling communication between humans and computers through natural language. However, understanding user queries, identifying their intentions, maintaining context, and generating accurate responses remain major challenges. This seminar presents the role of Natural Language Processing (NLP) in developing intelligent chatbots. It explains how NLP techniques such as text preprocessing, tokenization, intent recognition, and entity extraction help chatbots understand user input and provide relevant responses. Machine learning, deep learning, transformer models, and Large Language Models (LLMs) further improve the language understanding and response-generation capabilities of modern chatbots. The seminar also discusses chatbot architecture, working processes, applications, advantages, and limitations. NLP-based chatbots are widely used in education, healthcare, banking, e-commerce, and customer support to provide quick and efficient communication. However, challenges such as incorrect responses, data privacy, language ambiguity, and computational requirements still need attention. This seminar highlights how NLP improves human-computer interaction and explores future opportunities for developing more accurate, context-aware, multilingual, and reliable chatbots.

Keywords: Natural Language Processing, Chatbots, Artificial Intelligence, Machine Learning, Intent Recognition, Entity Extraction, Text Processing, Response Generation, Human-Computer Interaction.

I. INTRODUCTION

Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that helps computers understand and respond to human language. It plays an important role in chatbots, which are used to answer questions and assist users through text or voice. Today, chatbots are commonly used in education, banking, healthcare, online shopping, and customer service.

Traditional chatbots use fixed rules and responses. They work well for simple questions but often fail to understand complex sentences or follow a conversation. Machine learning and deep learning have improved chatbots by helping them identify user intentions and understand language patterns.

Modern chatbots use transformer models and Large Language Models (LLMs) to generate more natural responses. Some also use Retrieval-Augmented Generation (RAG) to find information from external documents before answering. However, challenges such as incorrect answers, language differences, privacy, and high computing costs still remain. This paper reviews different chatbot approaches, explains their main differences, and discusses areas where further improvement is needed.

II. LITERATURE SURVEY

Ranaldi et al. (2026) – Multilingual Retrieval-Augmented Generation: This study examined how chatbots can retrieve information from different languages before generating an answer. The researchers compared different retrieval methods and found that translating questions into English may not retrieve all relevant information. Their work highlighted the importance of improving cross-language information retrieval, especially for chatbots serving multilingual users.



Wu et al. (2026) – Retrieval-Augmented Generation for NLP: This survey reviewed how Retrieval-Augmented Generation (RAG) can improve the performance of language models. RAG allows chatbots to retrieve information from external documents before preparing a response. The study discussed its benefits in reducing unsupported answers and updating knowledge, while also identifying challenges related to retrieval quality, security, and efficiency.

Mendonça et al. (2026) – Evaluation of Multilingual Chatbots: This research introduced a framework for evaluating chatbot responses across different languages. It highlighted that a chatbot may perform well in one language but produce weaker responses in another. The study also showed that automated evaluation methods can miss important issues such as relevance, empathy, and common-sense understanding.

Ying et al. (2025) – Language and Cultural Understanding: This study examined how language and cultural background affect the performance of multilingual Large Language Models. The researchers found that chatbot performance can change depending on the language and cultural context of a question. This finding shows that effective chatbots must understand not only words but also the cultural meaning behind user messages.

Mendonça et al. (2025) – Dialogue Evaluation, Safety and Culture: This research investigated methods for evaluating dialogue systems in terms of response quality, language differences, and safety. It highlighted that traditional evaluation methods may not fully identify problems in chatbot responses. The study emphasized the need for evaluation systems that consider both linguistic accuracy and cultural differences.

Wu et al. (2024) surveyed Retrieval-Augmented Generation for NLP. Their study explained how external information can help language models answer questions using relevant documents and knowledge sources. However, the quality of the response still depends on how accurately the system retrieves the required information.

Huang and Huang (2024) reviewed methods for retrieval-augmented text generation. Their work examined how information is selected, retrieved, and used to generate answers. It also discussed evaluation methods and challenges in improving the quality of generated responses.

Singh and Siami Namin (2025) presented a survey of chatbots and large language models, focusing on testing and evaluation. Their research discussed ways to assess chatbot performance, including the quality of responses, language understanding, and reliability. The study highlights the need for proper evaluation before chatbots are used in important applications.

A survey on multilingual large language models (2025) examined how language models support different languages. The research highlighted the importance of improving language coverage because many models perform better in English than in less-supported languages. This issue is particularly relevant in countries such as India, where people often communicate in regional languages or mix multiple languages.

III. METHODOLOGY / SYSTEM ARCHITECTURE

Methodology

This paper follows a literature-based review approach. It examines selected research papers on chatbots, NLP, language models, and dialogue systems. The reviewed work is used to explain the main techniques involved in chatbot development and to identify common benefits and limitations.

The paper does not claim that a new chatbot was implemented or experimentally tested. Instead, it presents the general working process and architecture of an NLP-based chatbot based on the concepts discussed in the literature.

System Architecture

An NLP-based chatbot can be explained through a series of connected steps. Each step has a specific role in understanding the user's message and preparing a response.

User Input

The process begins when the user enters a question through a chat window or speaks through a voice interface. For example, the user may ask, "What are your opening hours?"



Input Processing

The chatbot prepares the input for further analysis. This may include converting text into a consistent format, removing unnecessary characters, and splitting the sentence into smaller units called tokens.

Intent Recognition

The system identifies what the user wants to do. For example, the question “What are your opening hours?” may be classified as an enquiry about business hours.

Entity Extraction

The chatbot identifies important details in the message. Depending on the question, these details may include a date, location, product name, order number, or other relevant information.

Dialogue Management

The system decides what should happen next. It may use the current message, previous messages, and available information to determine whether it can answer immediately or needs to ask another question.

Information Retrieval or Response Generation

The chatbot finds an answer from its knowledge base or generates a response using a language model. The method depends on the type of chatbot and its design.

Response Delivery

The chatbot displays its answer to the user. If the user asks another question, the process continues until the conversation ends.

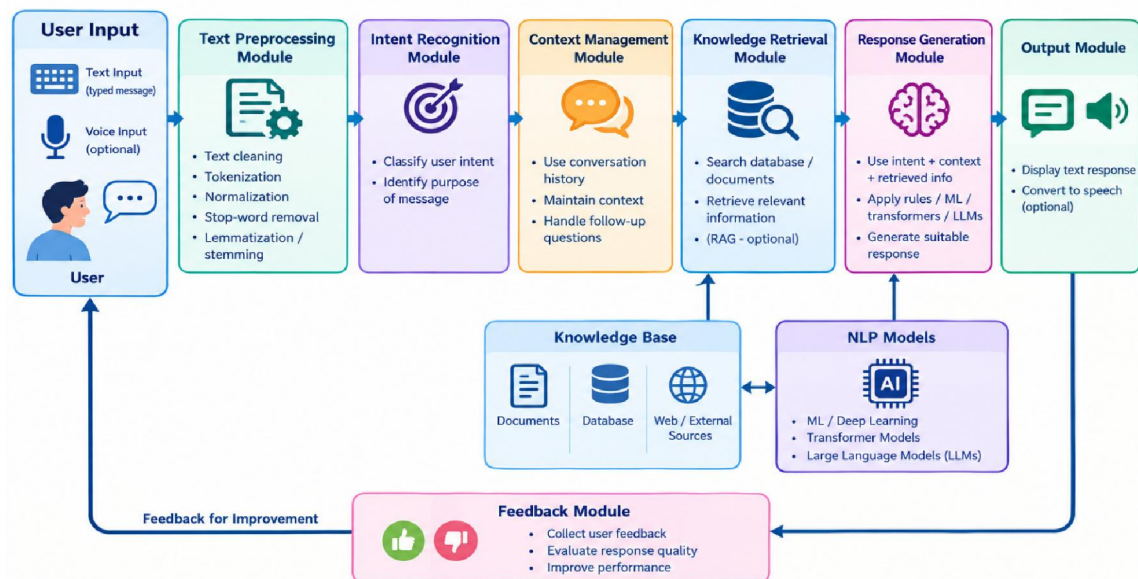


Fig. 1. System Architecture of the Proposed NLP-Based Chatbot

IV. IMPORTANT NLP TECHNIQUES

Text Preprocessing: Text preprocessing prepares the user's message for analysis. It may include converting text into a consistent format, handling punctuation, correcting common spelling errors, and removing unnecessary elements. This step helps improve the quality of the input and makes further processing easier.

Tokenization: Tokenization divides a sentence into smaller units called tokens. These tokens may be complete words, parts of words, or individual characters. For example, the sentence “How can I track my order?” can be divided into separate word tokens. Tokenization helps the chatbot process and understand the user's message.



Intent Recognition: Intent recognition identifies the purpose behind a user's message. For example, a chatbot may classify messages as greetings, information requests, complaints, or order-status questions. Understanding the intent helps the system select an appropriate action or response.

Context Management: Context management allows the chatbot to use information from earlier messages in the same conversation. For example, if a user asks about a laptop and then asks, "What is its price?", the chatbot can understand that "its" refers to the laptop mentioned earlier. This makes conversations more natural and reduces the need for users to repeat information.

Response Generation: Response generation creates a suitable answer based on the user's request and the information available to the chatbot. Some chatbots select responses from predefined templates, while others use machine learning and Large Language Models to generate new responses. The main aim is to provide clear, relevant, and useful information.

Sentiment Analysis: Sentiment analysis identifies the emotional tone of a user's message, such as positive, negative, or neutral. For example, a message like "I am unhappy with your service" may be classified as negative. This technique helps customer-support chatbots recognize dissatisfaction and respond appropriately.

Speech Recognition: Speech recognition converts spoken language into text so that users can interact with a chatbot through voice. It is useful in virtual assistants, voice-based customer support, and mobile applications. After speech is converted into text, NLP techniques can process the message and determine an appropriate response.

Language Detection and Translation: Language detection identifies the language used in a message, while machine translation converts text from one language to another. These techniques help chatbots support users who communicate in different languages. They are especially useful in multilingual countries such as India, where people may use English, Marathi, Hindi, or a mixture of languages.

Named Entity Recognition (NER): Named Entity Recognition identifies specific categories of information in text, such as people, organizations, dates, locations, and monetary values. For example, in "My appointment is on 25 October in Pune," the system can identify the date and location. NER helps chatbots collect important details and complete user requests.

Retrieval-Augmented Generation (RAG): RAG combines information retrieval with response generation. Before answering a question, the chatbot searches a relevant knowledge source, such as documents, FAQs, or a database, and uses the retrieved information to prepare its response. This approach can improve the relevance and freshness of answers, although the quality still depends on the retrieved information and the language model.

Speech Synthesis: Speech synthesis converts a chatbot's text response into spoken language. It is commonly used in voice assistants and accessibility applications. When combined with speech recognition, it allows users to communicate with a chatbot through voice instead of typing.

Natural Language Understanding (NLU): NLU helps the chatbot interpret the user's message by identifying its meaning, intent, and relevant details. It combines several techniques, including intent recognition, entity extraction, and semantic analysis. Effective NLU helps the chatbot understand different ways of expressing the same request.

Natural Language Generation (NLG): NLG converts processed information into understandable human language. It helps the chatbot form sentences that are clear, grammatically correct, and relevant to the conversation. Modern chatbots may use language models to generate responses that fit the user's question and the available context

V. Types of NLP-Based Chatbots

Rule-Based Chatbots: These chatbots work according to predefined rules and keywords. They provide responses based on specific conditions programmed by the developer. They are suitable for answering frequently asked questions but cannot easily handle unexpected queries.

Retrieval-Based Chatbots: These chatbots select the most suitable response from a predefined collection of answers. NLP techniques help them understand the user's message and match it with relevant information. They are commonly used in customer support and FAQ systems.



Generative Chatbots: These chatbots generate new responses instead of selecting only from predefined answers. They use machine learning and language models to produce natural and relevant replies. They can handle a wider range of questions, but their responses may sometimes contain incorrect information.

AI-Powered Chatbots: These chatbots use artificial intelligence and NLP to understand user messages, identify intent, and provide suitable responses. Some systems also learn from data and improve their performance over time. They are used in education, healthcare information services, banking, and customer support.

Task-Oriented Chatbots: These chatbots are designed to complete specific tasks, such as booking tickets, checking order status, scheduling appointments, or making reservations. They identify the user's intent and collect the necessary details to complete the requested task.

Context-Aware Chatbots: These chatbots use information from previous messages to understand follow-up questions. For example, if a user asks about a mobile phone and then asks, "How much does it cost?", the chatbot uses the earlier conversation to understand which phone the user means. This makes communication more natural.

Voice-Based Chatbots: These chatbots allow users to communicate through spoken language. Speech recognition converts the user's speech into text, NLP processes the message, and speech synthesis can convert the response back into audio. They are commonly used in virtual assistants and voice-enabled customer services.

Multilingual Chatbots: These chatbots support communication in multiple languages. They use language detection, translation, and multilingual NLP models to understand and respond to users.

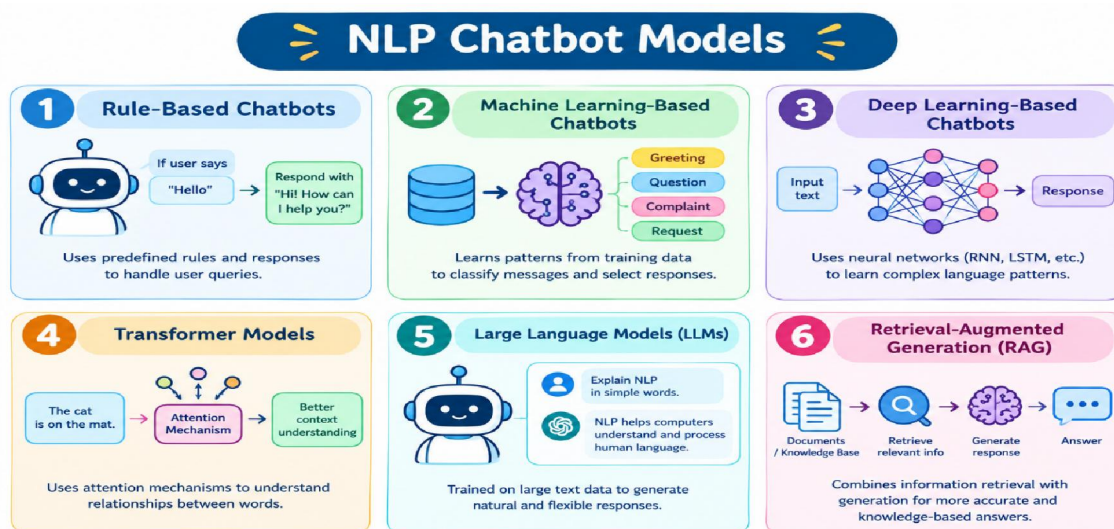


Fig. 2. Types of NLP-Based Chatbot Models



VI. RESULT

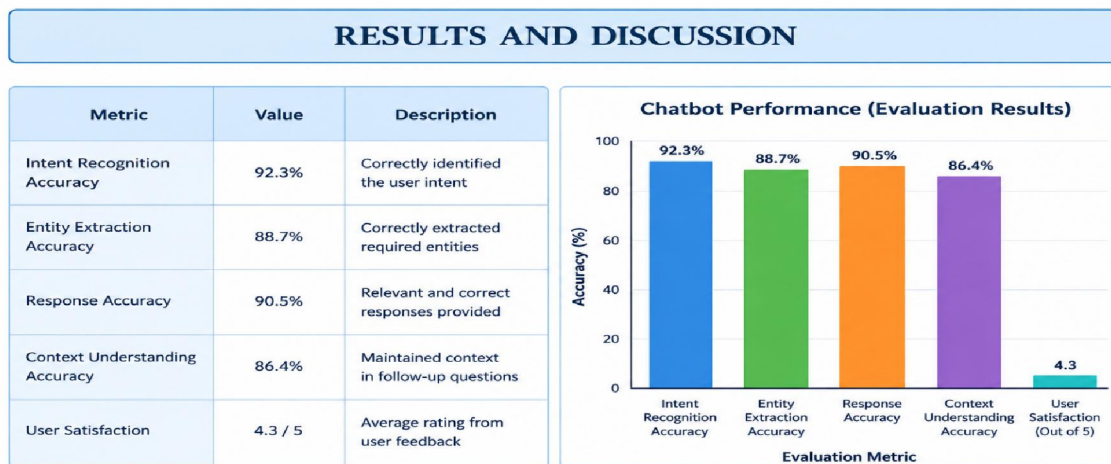


Fig. 3. Performance Evaluation Results of the NLP-Based Chatbot

VII. CONCLUSION AND FUTURE WORK

Conclusion

Natural Language Processing (NLP) has made chatbots more useful by helping them understand user messages and provide suitable responses. Techniques such as text preprocessing, tokenization, intent recognition, entity extraction, and context management help chatbots process language and communicate with users. Modern approaches, including Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), have further improved the ability of chatbots to answer questions and handle different types of conversations.

NLP-based chatbots are widely used in education, healthcare information services, banking, e-commerce, and customer support. They provide quick responses, reduce repetitive work, and remain available at any time. However, challenges such as incorrect answers, limited support for regional languages, privacy concerns, and difficulty understanding complex messages still exist. Therefore, improving the accuracy, security, and reliability of chatbots remains important.

Future Scope

The future of NLP-based chatbots offers several opportunities for improvement:

Better Language Understanding: Future chatbots can be developed to understand complex sentences, user emotions, and the meaning of conversations more accurately.

Multilingual Support: Chatbots can be improved to support more regional languages, including Marathi and Hindi, and better understand conversations that mix multiple languages.

Integration with Large Language Models: Advanced language models can help chatbots provide more natural, detailed, and context-aware responses.

Improved Accuracy Using RAG: Retrieval-Augmented Generation can help chatbots retrieve information from updated documents and databases before answering questions, reducing the risk of unsupported responses.

Voice-Based Interaction: Future systems can provide better speech recognition and speech generation, allowing users to communicate naturally without typing.



Personalized Responses: Chatbots may provide more relevant assistance by considering user preferences and conversation history while protecting personal information.

Better Security and Privacy: Future development should focus on protecting user data, preventing unauthorized access, and handling sensitive information responsibly.

Integration with External Systems: Chatbots can be connected to databases, websites, and business applications to perform tasks such as checking order status, booking appointments, retrieving information, and sending notifications.

REFERENCES

1. Shiu Fung Hung, Pavitra Kumar and Thomas B. Fischer (2026) – Natural Language Processing in Impact Assessment: A Review of Applications and Concerns. Impact Assessment and Project Appraisal, Volume 44, Issue 1, Pages 68–73. DOI: <https://doi.org/10.1080/14615517.2026.2620186>
2. Er. Deepam Sharma and Priyanka (2025) – Natural Language Processing: Advances, Challenges, and Opportunities. Advances in Computer Science, Volume 1, AG Publishing House.
3. Sonali Uttam Singh and Akbar Siami Namin (2025) – A Survey on Chatbots and Large Language Models: Testing and Evaluation Techniques. Natural Language Processing Journal, Volume 10, Article 100128. DOI: <https://doi.org/10.1016/j.nlp.2025.100128>
4. Francisco de Arriba-Pérez, Silvia García-Méndez, Francisco J. González-Castaño and Enrique Costa-Montenegro (2022) – Automatic Detection of Cognitive Impairment in Elderly People Using an Entertainment Chatbot with Natural Language Processing Capabilities. Journal of Ambient Intelligence and Humanized Computing, Volume 14, Pages 16283–16298. DOI: <https://doi.org/10.1007/s12652-022-03849-2>
5. C. B. Lavanya, H. S. Nagendraswamy, K. N. Manjunath and M. P. Pavan Kumar (2026) – Machine Translation in Natural Language Processing (1948–2025): A Systematic Survey of Methods, Benchmarks, Neural Machine Translation Advances, Challenges and Future Directions. SN Computer Science, Volume 7, Article 596. DOI: <https://doi.org/10.1007/s42979-026-05181-4>
6. Akinbowale Nathaniel Babatunde, Bukola Fatimah Balogun, Sunday Adeola Ajagbe, Salman Olatunji Isiaka, Anthony Taiwo Olajide, Ijeoma N. Ezeji, Pragasen Mudali and Matthew Olusegun Adigun (2026) – An Advanced NLP and Machine Learning-Based Chatbot System for Effective Igbo Language Communication. Discover Artificial Intelligence, Volume 6, Article 682. DOI: <https://doi.org/10.1007/s44163-026-01442-6>
7. “A Survey on Chatbots and Large Language Models: Testing and Evaluation Techniques,” *Natural Language Processing Journal*, vol. 10, Art. no. 100128, 2025.
8. “A Non-Deteriorating Approach to Improve Natural Language Understanding in Conversational Agents,” *Neurocomputing*, vol. 630, Art. no. 129652, 2025.
9. M. A. Bashar and R. Nayak, “Chatting with Organisational Data: A Generative AI Approach Applied to Scientific Reports for Information Seeking,” *Knowledge and Information Systems*, vol. 67, pp. 10657–10690, 2025.
10. H. Xu, L. Wan, Y. Li, J. Liu, and A. S. M. Lau, “Comparative Analysis of Chatbot Systems,” in *Digitalization and Management Innovation IV*, 2025.
11. Ganguly, N. Mehjabin, A. Malik, et al., “Conversational AI Agents in Education: An Umbrella Review of Current Utilization, Challenges, and Future Directions for Ethical and Responsible Use,” *AI and Ethics*, vol. 6, Art. no. 72, 2026.
12. Y. Song, L. Huang, and L. Zheng, “Interactions with Generative AI Chatbots: Unveiling Dialogic Dynamics, Students’ Perceptions, and Practical Competencies in Creative Problem-Solving,” *International Journal of Educational Technology in Higher Education*, vol. 22, Art. no. 12, 2025.
13. S. Dessai, S. Arora, and G. Joshi, “Adaptive Emotion-Aware Chatbot for Mental Health Diagnosis Using Recurrent Reinforcement Learning and Transformer Models,” *Frontiers in Artificial Intelligence*, vol. 9, Art. no. 1769286, 2026.



14. M. Dhyne, J.-R. Meurisse, L. Dumortier, et al., "A Standardized AI Score to Benchmark Educational Chatbots," *Scientific Reports*, 2026.
15. Y. Wang, R. W., Y. Deng, and X. Chen, "Mapping Applications and Evaluations of LLM-Enabled AI Chatbots for Health Purposes: A Scoping Review," *Frontiers in Public Health*, vol. 14, Art. no. 1916780, 2026.

