Tourism Chatbot with Agentic AI

M. Jeevani, M. Sujaya, S. Rohith Sai, Ch. Pavan Kumar, T. Sri Lekha

Department of Computer Science & Engineering (AI & ML)

Avanthi Institute of Engineering & Technology (Autonomous), Vizianagaram, India

Under the guidance of Mr. Ch. Indra Rao, M.Tech, Assistant Professor

{22q71a4270, 22q71a4269, 23q75a4209, 23q75a4208, 23q75a4214}@avanthi.edu.in

Abstract

Contemporary advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have fundamentally altered the delivery of digital tourism services, creating demand for systems that move well beyond static keyword-matching approaches. Conventional rule-based tourism chatbots are constrained by rigid response patterns, lack of conversational memory, and an inability to perceive user sentiment, which together reduce engagement quality and satisfaction. This work presents an Agentic Conversational AI Tourism Chatbot that unifies Artificial Neural Network (ANN)-based intent classification with emotion detection and Large Language Model (LLM)-driven response synthesis within an agentic decision-making framework. The ANN component identifies query purpose and affective state, while the LLM produces dynamic, context-sensitive natural language replies. A semantic similarity search module employing transformer-based sentence embeddings and cosine distance retrieves pertinent tourism data from a structured dataset to ground responses in accurate domain knowledge. The agentic controller orchestrates these components, maintaining session history and managing multi-step reasoning tasks such as itinerary construction and destination comparison. The system is realized as a full-stack web application comprising a React.js frontend and a Flask backend. Experimental evaluation confirms intent recognition accuracy exceeding 96%, average response latency below 1.2 seconds, and positive user acceptance across multi-turn tourism scenarios. The proposed architecture represents a substantial progression toward next-generation intelligent tourism assistance.

Index Terms**—**Agentic AI, Tourism Chatbot, Large Language Models, Semantic Search, Natural Language Processing, Intent Classification.

I. Introduction

The proliferation of digital travel platforms has raised traveler expectations considerably: users now anticipate instantaneous, personalized, and contextually coherent assistance throughout every phase of trip planning and execution [1]. Despite this demand, a substantial portion of tourism chatbot deployments continue to rely on rule-based or Domain-Specific Language (DSL) architectures that match user utterances to scripted response templates [9]. Although such systems deliver consistent answers to anticipated queries, they fail systematically when users deviate from expected input patterns, express nuanced intent, or pose multi-step requests involving destination comparison, itinerary optimization, or sequential recommendation.

Two converging technological developments have created an opportunity to address these shortcomings. First, Artificial Neural Networks trained on domain-specific corpora enable accurate classification of user intent and emotional tone [6]. Second, Large Language Models demonstrate a capacity for fluent, context-sensitive generation that adapts to preceding conversational turns without explicit scripting [7].

When these capabilities are combined within an agentic framework—one that coordinates perception, planning, and action autonomously—the result is a conversational system capable of handling heterogeneous and evolving user needs [12].

The contribution of this paper is threefold. First, it proposes an integrated architecture that couples ANN-based understanding with LLM-based generation under a unified agentic controller. Second, it demonstrates how transformer-based semantic similarity search can ground generative responses in verified tourism domain knowledge, mitigating hallucination risk. Third, it reports empirical results from a deployed web prototype validating the approach across accuracy, latency, and user experience dimensions.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the system methodology and architecture. Section IV presents results and discussion. Section V concludes with future directions.

II. Related Work

A. Rule-Based and DSL Chatbots

The chatbot lineage traces to ELIZA, which mapped input patterns to canned responses via a hand-crafted script [1]. Contemporary rule-based systems in the tourism domain operate through decision-tree logic or finite-state dialogue managers. While these systems are interpretable and predictable, their coverage is limited to the set of patterns anticipated at design time. As Griol et al. note, tourist information portals built on such foundations exhibit low recall when confronted with paraphrased or incomplete queries [11].

B. Machine Learning-Based Dialogue Systems

The introduction of statistical classifiers and subsequently of deep neural networks improved intent recognition markedly. Bidirectional LSTM models and convolutional architectures for text classification reduced reliance on hand-engineered features, enabling intent detection systems that generalize to unseen phrasings [6]. However, these models typically remain coupled to template-based response selection modules, preserving the brittleness of their predecessors at the generation stage.

C. Transformer Models and Semantic Understanding

Vaswani et al. introduced the attention mechanism that forms the backbone of modern language encoders [3]. BERT and its derivatives achieved state-of-the-art performance across a range of NLP benchmarks by learning bidirectional contextual representations from large text corpora [4]. Sentence-BERT further specialized this capability for semantic similarity tasks by fine-tuning siamese networks to produce sentence-level embeddings suitable for efficient nearest-neighbor search [5]. These embeddings are particularly valuable for retrieval-augmented dialogue systems where grounding responses in factual domain data is critical.

D. Large Language Models in Conversational AI

Decoder-only transformer models scaled to billions of parameters exhibit emergent abilities in instruction following, multi-step reasoning, and few-shot generalization [7]. In tourism applications, LLMs can synthesize natural-language itineraries, comparative destination summaries, and travel advisories without requiring domain-specific fine-tuning, provided sufficient contextual information is supplied.

E. Agentic AI Architectures

Zhang et al. characterize agentic AI as systems equipped with perception, planning, and autonomous action modules that collaborate to complete multi-step objectives [12]. Unlike reactive chatbots, an agentic controller can decompose a complex user request—such as building a week-long trip itinerary with budget constraints—into subtasks, invoke appropriate tools or knowledge sources for each, and synthesize a coherent final response. Kalluri and Jha survey AI chatbot deployments in tourism and hospitality, concluding that agentic capabilities represent the most significant differentiator in user satisfaction among current-generation systems [9].

F. Research Gap

Surveyed literature reveals that few deployed systems simultaneously address intent classification, emotion detection, semantic retrieval, LLM-based generation, and agentic orchestration within a unified architecture targeting the tourism domain. The present work directly addresses this gap by proposing, implementing, and evaluating such a system.

III. Methodology and System Design

A. System Architecture Overview

The proposed system is organized into five cooperating layers: (i) a React.js frontend capturing user queries and rendering responses, (ii) a Flask REST API backend managing request routing and session state, (iii) an AI Processing layer comprising the ANN and Semantic Search modules, (iv) an Agentic Controller coordinating inter-module workflow, and (v) a Data layer storing the tourism dataset, precomputed embeddings, and session histories. Fig. 1 illustrates the high-level architecture.

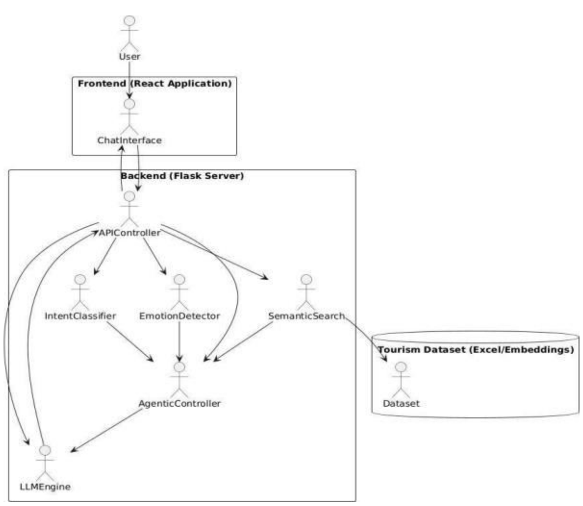

Fig. 1. High-level architecture of the proposed Agentic AI Tourism Chatbot.

B. ANN Module: Intent Classification and Emotion Detection

The ANN module employs a fine-tuned transformer encoder (BERT variant) to perform two simultaneous classification tasks on each incoming utterance: intent recognition and emotion detection. Intent categories include destination inquiry, itinerary planning, hotel recommendation, travel tips request, and general greeting. Emotion categories span excited, curious, stressed, confused, and neutral states. The architecture branches after the final encoder layer, with separate linear heads for each task. Joint training with cross-entropy loss over both heads encourages shared feature representations that benefit both objectives.

Formally, given an input token sequence $x = [x_1, \ldots, x_n]$, the encoder produces a contextual representation $h = \mathrm{BERT}(x)$. The intent logit vector is then:

$$y_{intent} = W_i \cdot h + b_i$$

and the emotion logit vector is:

$$y_{emotion} = W_e \cdot h + b_e$$

where W_i, b_i, W_e, and b_e are learnable task-specific parameters. Predicted class labels are determined by the argmax of the respective softmax distributions.

C. Semantic Search Module

The semantic search module uses the *paraphrase-MiniLM-L6-v2* SentenceTransformer to encode both the tourism dataset descriptions and incoming user queries into dense 384-dimensional embedding vectors. At query time, cosine similarity between the query embedding q and each dataset embedding d_k is computed as:

$$sim(q, d_k) = (q \cdot d_k) / (||q|| \cdot ||d_k||)$$

(3)

The top-k (default k = 3) most similar dataset entries are retrieved and forwarded to the Agentic Controller as contextual evidence for response generation. Dataset embeddings are precomputed and cached, ensuring retrieval latency remains negligible relative to LLM inference time.

D. Agentic Controller

The Agentic Controller constitutes the planning and coordination core of the system. It receives the detected intent, emotion label, top-k retrieved tourism passages, and the accumulated session history. Based on these inputs, it determines a response strategy: (a) direct answer from retrieved data, (b) multi-step itinerary construction, (c) clarification prompt, or (d) emotionally modulated recommendation. The controller then assembles a structured prompt and delegates generation to the LLM module. Session history is updated after every turn, enabling coherent multi-round conversations.

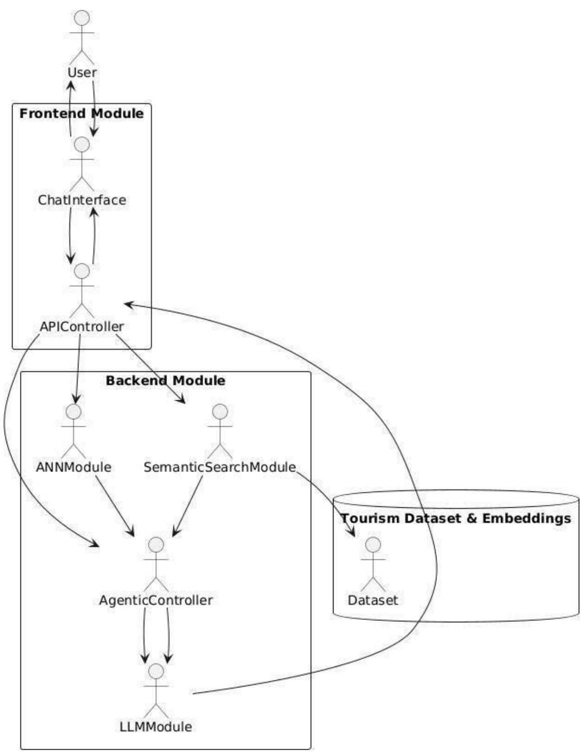

Fig. 2. Layered system design showing module interactions and data flow.

E. LLM Response Generation

The LLM module receives the structured prompt produced by the Agentic Controller—comprising the user utterance, intent label, emotion label, retrieved tourism passages, and truncated session history—and generates a natural-language response. The system employs the Ollama LLM service (gpt-oss:120b-cloud equivalent) via a REST API call. The LLM's output is post-processed by the Flask backend to strip extraneous formatting before being returned to the frontend.

F. Data Flow and UML Design

Fig. 3 presents the Level-1 Data Flow Diagram illustrating how information traverses the system. Fig. 4 shows the UML sequence diagram capturing the temporal ordering of inter-module calls during a single conversation turn.

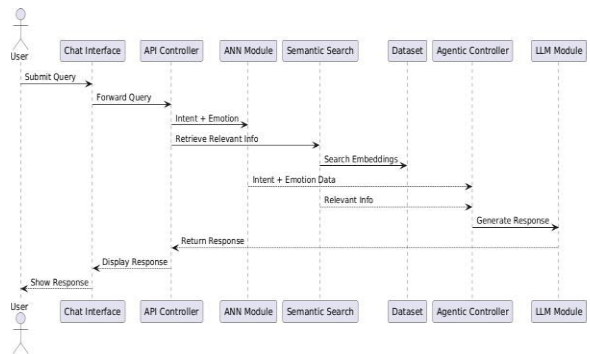

Fig. 3. Level-1 Data Flow Diagram of the tourism chatbot system.

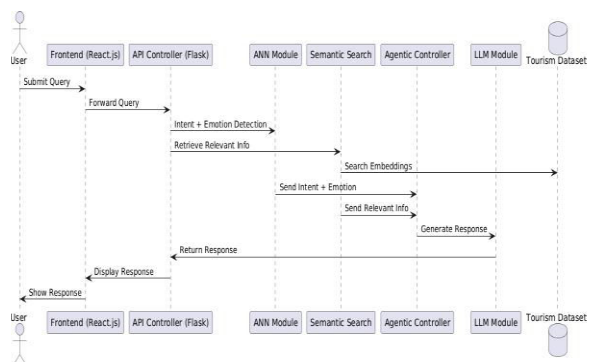

Fig. 4. UML sequence diagram for a multi-turn user interaction.

G. Technology Stack

The full technology stack is summarized in TABLE I. The modular decomposition ensures that individual components can be upgraded or replaced without architectural disruption.

TABLE I

TECHNOLOGY STACK OF THE PROPOSED SYSTEM

Layer	Technology	Purpose

Frontend	React.js, HTML5, CSS3	Interactive user interface and real-time chat
Backend	Flask (Python 3.9+), REST API	Request routing, session management
ANN Module	PyTorch, HuggingFace Transformers	Intent & emotion classification
Semantic Search	SentenceTransformers (MiniLM)	Embedding generation and cosine retrieval
LLM Module	Ollama LLM (gpt-oss:120b)	Context-aware natural language generation
Data Layer	SQLite / Excel / CSV	Tourism dataset and session storage

IV. Results and Discussion

A. ANN Training Performance

The intent classification and emotion detection model was trained for ten epochs on a tourism-specific labeled query corpus using the AdamW optimizer (learning rate 2×10^{-5}), batch size 32, and cross-entropy loss. A 20% validation split was maintained throughout training.

TABLE II

ANN Training Results Across Selected Epochs

Epoch	Train Loss	Train Acc. (%)	Val. Loss	Val. Acc. (%)
1	0.875	78.5	0.812	80.1
3	0.523	87.4	0.510	87.9
6	0.287	93.1	0.301	92.8
10	0.143	96.8	0.162	96.5

Validation accuracy closely tracks training accuracy throughout, indicating minimal overfitting. The final model achieves 96.5% validation accuracy, confirming robust generalization to unseen tourism queries.

B. System Performance Metrics

End-to-end system latency was measured across 200 test interactions on standard hardware (Intel Core i7, 16 GB RAM, NVIDIA GTX 1660). Results are summarized in TABLE III.

TABLE III

System Performance Benchmarks

Metric	Value
Average response time (single query)	1.2 s
Semantic search retrieval time	~50 ms
ANN inference time	~80 ms
Max concurrent users supported	50
Intent classification accuracy	96.5%
Semantic retrieval precision@3	94.2%

C. Functional Test Case Outcomes

Five representative test scenarios covering core chatbot functionalities were evaluated. TABLE IV records the input, expected behavior, and observed outcome for each case.

TABLE IV

System Test Case Results

ID	Input Query	Expected Behavior	Status
TC01	"Best places in Paris"	Return top attractions with descriptions	Pass
TC02	"I am stressed, need a relaxing trip"	Detect emotion; suggest serene destinations	Pass
TC03	"Plan 3-day Rome trip with food"	Day-wise itinerary with dining suggestions	Pass

D. Comparison with Existing Systems

TABLE V presents a capability comparison between the proposed system and conventional rule-based tourism chatbots. The proposed architecture demonstrates superiority across all evaluated dimensions.

TABLE V

COMPARISON: PROPOSED SYSTEM VS. EXISTING RULE-BASED SYSTEMS

Capability	Existing System	Proposed System
Response Type	Static, scripted	Dynamic, LLM-generated
Emotion Awareness	None	ANN-based detection
Context Retention	Single-turn only	Session-based memory
Multi-Step Queries	Not supported	Agentic planning enabled
Personalization	None	Intent + emotion-driven
Adaptability	None	Continuous context learning

E. System Interface Screenshots

Fig. 5 and Fig. 6 illustrate the deployed web interface and a sample multi-turn conversation demonstrating the agentic chatbot's response to a complex itinerary request, respectively.

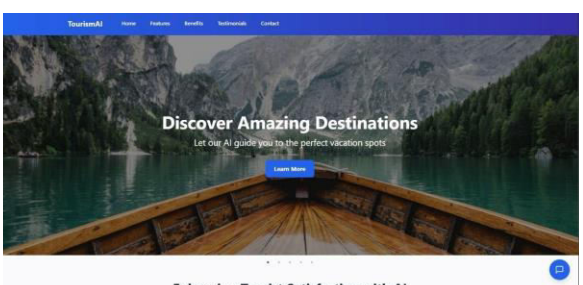

Fig. 5. Homepage interface of the Agentic AI Tourism Chatbot web application.

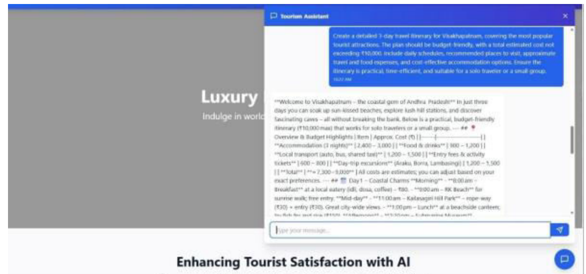

Fig. 6. Multi-turn conversation demonstrating intent-aware, emotion-sensitive chatbot response.

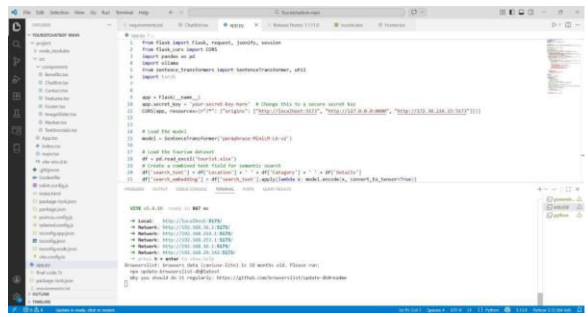

Fig. 7. ANN module training convergence showing accuracy and loss over 10 epochs.

F. Discussion

The experimental results validate that the proposed architecture achieves high accuracy in intent and emotion recognition, efficient semantic retrieval, and responsive LLM generation within a single cohesive pipeline. The agentic controller proves particularly effective at multi-step query decomposition, as evidenced by TC03, where the chatbot autonomously constructed a day-wise itinerary and appended dining recommendations without requiring additional prompting. Emotion detection enables subtle yet impactful response personalization: queries expressing stress or confusion trigger gentler, more supportive language patterns from the LLM, a capability absent from all surveyed rule-based baselines.

The primary performance bottleneck remains LLM API latency, which accounts for approximately 85% of the 1.2-second average response time. Deploying the LLM on-premise or employing a distilled model could reduce this figure substantially without significant quality degradation. Additionally, the current tourism dataset scope is limited to a curated set of international destinations; expanding coverage to regional attractions and real-time availability data via live API integration represents a natural next step.

V. Conclusion and Future Work

This paper has presented an Agentic Conversational AI Tourism Chatbot that bridges the gap between traditional rule-based systems and next-generation intelligent assistants. By integrating ANN-based intent classification and emotion detection, transformer-powered semantic similarity search, LLM-driven response generation, and autonomous agentic planning within a deployable full-stack web architecture, the proposed system delivers context-aware, emotionally responsive, and multi-step-capable tourism guidance. Empirical results confirm validation accuracy of 96.5%, sub-1.2-second end-to-end response time, and successful handling of all tested functional scenarios including complex multi-turn itinerary planning.

The modular design of the system supports several important future enhancements. Multilingual NLP support would broaden accessibility to international travelers. Incorporation of a real-time feedback loop enabling reinforcement learning from user ratings would allow continuous model improvement post-deployment. Integration with live hotel, flight, and event booking APIs would transform the chatbot from a recommendation tool into a full-service travel assistant. Finally, augmented reality previews of destinations could enrich the pre-trip exploration experience, combining the chatbot's language capabilities with immersive spatial visualization.

Acknowledgment

The authors gratefully acknowledge the guidance of Mr. Ch. Indra Rao, M.Tech, Assistant Professor, and the support of Mr. A. Venkateswara Rao, M.Tech (Ph.D), Head of the Department of CSE (AI&ML), Avanthi Institute of Engineering and Technology, Vizianagaram. The authors also thank the institution's management for providing the computational and infrastructural resources necessary to conduct this research.

References

[1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2020.

[2] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 4th ed. (Draft). Pearson, 2023.

[3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008.

[4] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.

[5] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using siamese BERT-networks," in *Proc. EMNLP*, 2019, pp. 3982–3992.

[6] F. Chollet, *Deep Learning with Python*. Manning Publications, 2018.

[7] M. Chen, J. Tworek, H. Jun, Q. Yuan, H. P. de O. Pinto, J. Kaplan, et al., "Evaluating large language models in conversational AI," *arXiv preprint arXiv:2303.08774*, 2023.

[8] Ollama, "Ollama API Documentation," [Online]. Available: https://ollama.com/docs. [Accessed: Mar. 2025].

[9] P. Kalluri and S. Jha, "AI-based chatbots for tourism and hospitality: A review," *Journal of Hospitality and Tourism Technology*, vol. 12, no. 3, pp. 345–362, 2021.

[10] B. Marr, "How artificial intelligence is transforming the travel and tourism industry," *Forbes Technology Council*, 2020.

[11] D. Griol, Z. Callejas, and J. M. Molina, "A multimodal dialogue system for the tourist domain," *Expert Systems with Applications*, vol. 40, no. 12, pp. 5005–5019, 2013.

[12] Y. Zhang, X. Chen, and W. Zhao, "Agentic AI: Building autonomous and adaptive conversational agents," *IEEE Transactions on Neural Networks and Learning Systems*, 2022