**FACULTY OF ENGINEERING, SCIENCE AND**

**TECHNOLOGY**

**SCHOOL OF COMPUTING**

**BACHELOR OF INFORMATION TECHNOLOGY(HONS)**

**Final Year Project-I**

**(EC3319)**

**Chatbot for Tourists Guide in Nepal Using NLP**

Name: Dinesh Bista

Student ID: 00020517

Supervisor: Romesh Poudyal

Submission Date: 13th May 20

This project is submitted in fulfilment of the requirements for Final Year Project-I of BIT (Hons), Nilai University.

**SUPERVISOR DECLERATION**

I hereby declare that I have checked this project and, in my opinion, this project will adequate in terms of scope and quality for the award of the degree of Bachelor of Information Technology.

**Signature:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Supervisor:** Er. Romesh Poudyal

**Date:** 12th May, 2025

**STUDENT’S DECLARATION**

I declare that this report entitled “**Chatbot for tourist guid in Nepal using NLP**” is my work. I have appropriately cited all sources. The report has not been accepted for any diploma or degree and is not being submitted simultaneously in candidature for any degree or other award.

**Signature**: Dinesh

**Name of Supervisor:** Mr. Dinesh Bista

**Date:** 12th May, 2025

**Acknowledgement**

I would like to express my sincere gratitude to my project supervisor, Er. Ramesh Poudyal, for their invaluable guidance and support throughout the project titled "Chatbot for Tourist Guide in Nepal Using NLP.

Additionally, I extend my thanks to Padmashree College and Nilai University for providing the necessary resources and facilities that were essential for the completion of this project. I appreciate the contributions of all individuals and institutions that have directly or indirectly supported me in this endeavor. Your encouragement and support have been deeply appreciated.

**Abstract**

This project proposes the development and implementation of a chatbot that will serve as an interactive and intelligent tourist guide for Nepal, utilizing advanced Natural Language Processing (NLP) techniques. The primary objective will be to enhance the travel experience for both local and international tourists by offering comprehensive, real-time assistance and information. The chatbot will be designed to understand and process natural language queries, enabling tourists to interact with it conversationally and receive accurate and timely responses. It will include key features such as location guidance, providing detailed information about tourist destinations, directions, maps, and transportation options. Real-time updates on weather conditions, local events, and travel advisories will be offered to help tourists plan effectively. A tracking feature will allow users to record visited or planned locations, assisting with itinerary organization. Additionally, the chatbot will recommend suitable accommodations such as homestays and hotels based on user preferences and budget, with details on amenities, pricing, and availability. It will also suggest cultural experiences, adventure activities, and sightseeing options tailored to user interests. Multilingual support will ensure accessibility for tourists from diverse linguistic backgrounds. The chatbot will provide personalized recommendations for dining and exploring, creating a more tailored travel experience. In emergencies, it will deliver essential information on nearby hospitals, police stations, and embassies, including contact details. Ultimately, the “Chatbot for Tourist Guide in Nepal” will aim to become an indispensable digital companion, enhancing tourism through timely, accurate, and personalized support.

**Keywords**: Chatbot, Tourist, Location Guidance, Real-Time Updates, Accommodation Recommendations, Personalized Recommendations, Emergency Assistance.**List of Abbreviations:**

**AI:** Artificial Intelligence

**NLP**: Natural Language Processing

**API**: Application Programming Interface

**GPS**: Global Positioning System

**ML**: Machine Learning

**DL**: Deep Learning

**NLTK**: Natural Language Toolkit

**UI**: User Interface

**UX**: User Experience

**HTML**: Hypertext Markup Language

**CSS**: Cascading Style Sheets

**JS**: JavaScript

**DBMS**: Database Management System

**SDK**: Software Development Kit

**IDE**: Integrated Development Environment

Table of Contents

[1. Introduction 1](#_Toc198157921)

[2. Objectives 3](#_Toc198157922)

[3. Research Background 4](#_Toc198157923)

[3.1 Problem statement: 4](#_Toc198157924)

[3.2 Scope: 5](#_Toc198157925)

[4. Literature Review: 6](#_Toc198157926)

[4.1 Artificial Intelligence Markup Language (AIML) 6](#_Toc198157927)

[4.2 Dialogue State Tracking (DST) and Multi-Domain Chatbots 7](#_Toc198157928)

[4.3 NLP-Based Recommender Chatbots 8](#_Toc198157929)

[4.4 Large Language Models (LLMs) and TourLLM 9](#_Toc198157930)

[4.5 ChatGPT in e-Tourism 10](#_Toc198157931)

[4.6 Summary and Comparative Analysis of Literature 11](#_Toc198157932)

[5. Accuracy Comparison Table 11](#_Toc198157933)

[6. Comparative Analysis with Our Project 12](#_Toc198157934)

[7. Gaps in Literature and Relevance of This Study Summary 13](#_Toc198157935)

[8. Research Methodology 14](#_Toc198157936)

[8.1 Description of Methodology 14](#_Toc198157937)

[8.2 Materials and Equipment Used 15](#_Toc198157938)

[8.3 Dataset Collection and Sample Preparation 16](#_Toc198157939)

[8.4 Sampling and Randomization Techniques 17](#_Toc198157940)

[8.5 Measurements and Calculations 17](#_Toc198157941)

[8.6 Statistical Analysis Techniques 18](#_Toc198157942)

[9. Schedule and Deliveries: 19](#_Toc198157943)

[9.1 Gantt Chart 19](#_Toc198157944)

[9.2 Milestones 19](#_Toc198157945)

[10. Expected Outcome 20](#_Toc198157946)

[11. Conclusion 21](#_Toc198157947)

[12. References 22](#_Toc198157948)

[13. Appendix 24](#_Toc198157949)

[13.1 Poster Presentation 24](#_Toc198157950)

[13.2 Project log file 25](#_Toc198157951)

**Table of Images**

[Figure 1 : Chatbot for Tourist Guide in Nepal 1](file:///D:\E%20drive\Users\Acer\Desktop\Cap%201%20(AutoRecovered).docx#_Toc198157972)

[Figure 2 : Workflow Diagram 14](#_Toc198157973)

[Figure 3 : Gantt Chart for project activities 19](#_Toc198157974)

[Figure 4 : Poster Presentation 24](file:///D:\E%20drive\Users\Acer\Desktop\Cap%201%20(AutoRecovered).docx#_Toc198157975)

**List of Tables**

[Table 1 : Comparative Analysis with project 21](#_Toc198072476)

[Table 2 : Milestones 28](#_Toc198072477)

# 1. Introduction

Tourism will continue to be a significant contributor to Nepal’s economy, playing a vital role in promoting cultural exchange, generating employment, and supporting local businesses. With its diverse natural landscapes, ancient heritage sites, and spiritual experiences, Nepal is expected to remain a popular destination for tourists from around the world. However, many visitors especially those unfamiliar with the country are likely to face difficulties in accessing real-time information, understanding the local language, and navigating locations efficiently. These challenges may affect their overall travel experience and limit their ability to explore Nepal’s full tourism potential.

Figure 1 : Chatbot for Tourist Guide in Nepal

In the coming years, the use of Artificial Intelligence (AI) is anticipated to grow substantially in the tourism and hospitality sectors. Among the most promising AI applications will be chatbots automated systems designed to simulate human-like conversations. These systems will offer instant responses, reduce reliance on human assistance, and operate around the clock. When integrated with Natural Language Processing (NLP), chatbots will become even more effective by understanding and responding to user queries in a natural and intuitive manner.

This project will focus on the development of a chatbot-based tourist guide specifically designed for Nepal, utilizing NLP techniques. The goal will be to create an intelligent system capable of understanding user inputs in both English and Nepali and providing helpful travel-related information such as places to visit, cultural insights, food recommendations, hotel locations, weather updates, and emergency contacts. The chatbot will also be equipped to answer frequently asked questions (FAQs) from tourists and offer personalized recommendations based on user preferences.

The implementation of this project will involve using Python as the primary programming language, along with frameworks such as Rasa or Dialog flow for developing the NLP model. Techniques like intent classification, named entity recognition (NER), and rule-based fallback handling will be employed to interpret the context of user queries and generate appropriate responses. The chatbot will be accessible through a web-based or mobile interface, ensuring convenience for tourists during their stay in Nepal.

By integrating AI and language technologies into Nepal’s tourism infrastructure, this project will aim to enhance the visitor experience, reduce communication barriers, and make travel information more readily available. It will also foster digital innovation in the local tourism industry by addressing long-standing challenges through modern solutions. In the future, the chatbot may be enhanced to support voice commands, integrate with live maps, and connect with booking systems ultimately evolving into a complete digital travel companion.

In conclusion, the “Chatbot for Tourist Guide in Nepal Using NLP” will represent a meaningful step toward transforming traditional tourism support systems into smart, scalable, and efficient digital assistants that improve service delivery, foster cultural understanding, and meet the evolving needs of modern travelers.

# 2. Objectives

* To develop a chatbot system using Natural Language Processing (NLP) that provides accurate and helpful travel-related information about Nepal to tourists.
* To enable the chatbot to understand and respond to user queries in both English and Nepali languages, enhancing accessibility for local and international travelers.

# 3. Research Background

## 3.1 Problem statement:

Tourism plays a vital role in the economic and cultural growth of many countries, including Nepal. The country, known for its rich history, cultural diversity, and natural beauty, attracts millions of tourists every year. However, despite the growth in tourist numbers, there is a noticeable gap in providing real-time, accessible, and multilingual information to travelers. Many tourists struggle with language barriers, insufficient information on transportation, accommodation, and attractions, which affects their travel experience. In this context, a chatbot powered by Natural Language Processing (NLP) can bridge the information gap by offering an interactive, personalized, and real-time communication tool for tourists, enhancing their journey through seamless assistance in their native language.

Tourists visiting Nepal often encounter difficulties in accessing comprehensive travel information due to language barriers and a lack of easily accessible, personalized resources. Traditional sources, such as travel guides and websites, may not provide the real-time or interactive support tourists need while traveling in a foreign country. For instance, a tourist in Kathmandu might have trouble finding the best route to a specific temple or understanding the local customs related to a particular festival. These gaps not only lead to frustration but also affect the overall travel experience, making it harder for tourists to fully enjoy their visit. Additionally, a significant number of travelers may not be familiar with Nepali, limiting their ability to communicate with locals and understand essential information.

The need for an interactive, real-time information system for tourists in Nepal is evident. A multilingual chatbot will address these issues by providing instant and relevant travel assistance through both text and voice. This system will offer personalized responses in multiple languages, such as English and Nepali, making it accessible to both international visitors and local tourists. For example, a chatbot will be able to answer questions like “What is the best way to get from Pokhara to Kathmandu?” Providing this information quickly and accurately will enhance the tourist experience, reduce confusion, and increase the likelihood of tourists recommending Nepal to others. Moreover, such a system will operate 24/7, making it an ideal solution for tourists who require assistance at any time of the day or night.

## 3.2 Scope:

This research will focus on the development of a chatbot system designed to provide essential travel information to tourists visiting Nepal. The chatbot will leverage Natural Language Processing (NLP) to understand and respond accurately to queries in English. The primary goal will be to create an intuitive and user-friendly platform that delivers real-time information, enhancing the tourist experience.

The scope of this research will be limited to the following aspects:

* Development of a chatbot system that will offer travel information in English.
* Utilization of Natural Language Processing (NLP) to effectively understand and respond to user queries.
* Provision of key travel information, including details about tourist destinations, accommodation, transportation, local customs, and emergency services.
* Design of an intuitive, user-friendly interface that will be accessible through mobile apps, websites, or messaging platforms.
* Limitation to responses in the English language for the chatbot.
* Exclusion of advanced machine learning techniques for continuous chatbot improvement, which may be explored in future research.

# 4. Literature Review:

The development of chatbot systems in tourism has evolved from simple rule-based agents to intelligent, learning-based conversational systems that utilize NLP and deep learning. This section reviews the primary algorithms and technologies explored in existing studies, with a focus on their application, strengths, limitations, and improvements—including future work suggested by researchers to enhance chatbot performance in real-world settings like Nepal.

## 4.1 Artificial Intelligence Markup Language (AIML)

AIML is a pattern-based, rule-driven markup language designed for creating conversational agents. It is particularly suitable for closed-domain systems where user queries are predictable and the scope is limited.

Ali et al. (2023) implemented an AIML-based chatbot to guide tourists in Sulaimani City. Their bot categorized questions into AIML files such as hotels, parks, restaurants, and museums, with over 500 categories and 352 Q&A pairs. The chatbot achieved 71.4% accuracy on in-domain queries and a user satisfaction score of 7.01 out of 10, demonstrating its usefulness for small-scale, domain-specific deployments.

**Benefits**:

* Simple and lightweight.
* Easy to deploy with minimal hardware.
* Transparent rule logic for specific information delivery.

**Weaknesses**:

* Cannot handle natural language variations.
* No learning or self-adaptation capability.
* Poor performance on cross-domain or unseen queries.

**Improvements Suggested / Future Work**:

* **Hybrid models**: Combine AIML with NLP or ML modules to handle ambiguous queries.
* **Voice input integration**: Enable audio-based interaction for hands-free use.
* **Auto-expansion**: Use data mining or user logs to automatically generate new AIML categories.
* **Deploy on real websites** and mobile apps to reach wider audiences.
* **Forward unknown queries** to administrators and allow the bot to learn from their responses.

## 4.2 Dialogue State Tracking (DST) and Multi-Domain Chatbots

Dialogue State Tracking (DST) forms the backbone of modern multi-turn conversational systems by preserving the user's goals, intents, and preferences across multiple dialogue turns.

Kim et al. (2023) proposed a multi-domain chatbot for tourism applications that incorporated DST for handling tasks such as hotel booking, finding food spots, and planning routes all within a single session. Their model managed seamless context switching and multi-intent dialogues, a significant improvement over rule-based systems.

**Benefits:**

* Retains user context across sessions, enabling fluid, goal-oriented conversations.
* Efficiently handles complex queries that span multiple subdomains.
* Offers dynamic suggestions based on accumulated dialogue history.

.

**Weaknesses:**

* Requires a substantial amount of labelled training data to identify intents and slots.
* High computational requirements, making it less feasible for small-scale developers.
* Increased complexity in system design and maintenance.

.

**Improvements Suggested / Future Work:**

* Incorporate pre-trained intent classification models to reduce manual effort.
* Extend support for multi-language environments, crucial for countries like Nepal.
* Add personalization layers through user profiling and history tracking.
* Use rule-based modules (like AIML) as fallbacks for specific FAQ-type queries.

## 4.3 NLP-Based Recommender Chatbots

Recommender chatbots use NLP techniques to extract user intent and provide personalized suggestions for destinations, itineraries, and experiences.

Mohite et al. (2024) presented a hybrid system that combined keyword extraction with rule-based filtering to tailor travel suggestions. The bot aimed to improve engagement by simplifying decision-making for tourists and offering contextual recommendations.

**Benefits:**

* + Improves user satisfaction by delivering relevant, personalized content.
  + Facilitates quick decisions by narrowing down choices based on preferences.
  + Allows real-time interaction with adaptive responses.

**Weaknesses:**

* Often lacks deeper context awareness, making it unsuitable for prolonged dialogue.
* Mostly static rule sets, which don’t learn from user interaction or feedback.
* Not well-suited for emotional or mood-based interpretation of queries.

**Improvements Suggested / Future Work**:

* + Use unsupervised ML techniques like clustering to segment user preferences.
  + Introduce feedback mechanisms where the chatbot learns from user ratings and selections.
  + Employ sentiment analysis to tailor suggestions according to the user's emotional tone.
  + Incorporate local travel datasets (e.g., Nepali destinations) for enhanced relevance.

## 4.4 Large Language Models (LLMs) and TourLLM

LLMs such as GPT-4 and domain-adapted versions like TourLLM are capable of generating human-like, context-rich responses across multiple domains.

Wei et al. (2024) presented TourLLM, an LLM fine-tuned on tourism data to improve accuracy and contextual understanding. Their study showed that domain-specific LLMs significantly outperform general-purpose models in the tourism context.

**Benefits**:

* + Handles natural, open-ended queries across various subtopics.
  + Supports multi-language and multi-domain functionality with coherent output.
  + Capable of learning and adapting through interaction, especially in online systems.

**Weaknesses**:

* + Computationally expensive and requires significant memory/storage resources.
  + May generate hallucinated or inaccurate information without proper grounding.
  + Output can be inconsistent without careful fine-tuning or validation layers.

**Improvements Suggested / Future Work**:

* Fine-tune with local tourism data (e.g., Nepal Tourism Board datasets).
* Combine with structured knowledge bases such as city maps, hotel APIs, and weather services.
* Implement Reinforcement Learning from Human Feedback (RLHF) to improve factual correctness.
* Add multilingual voice support for broader accessibility to diverse tourist populations.

## 4.5 ChatGPT in e-Tourism

Mich & Garigliano (2023) explored how ChatGPT can enhance e-tourism platforms by providing natural, conversational assistance for trip planning, hotel inquiries, and destination guidance. ChatGPT’s fluency and adaptability make it an attractive choice for tourism applications.

**Benefits**:

* Delivers highly fluent, human-like responses in multiple languages.
* Generates real-time, context-sensitive content based on user input.
* Easily integrated into websites, mobile apps, or messaging platforms.

**Weaknesses**:

* Cannot access or update real-time data (e.g., weather, hotel availability) unless connected to APIs.
* May occasionally generate factually incorrect or misleading information.
* Personalization remains limited unless specifically fine-tuned per user base.

**Improvements Suggested / Future Work**:

* ntegrate tourism-specific APIs (e.g., TripAdvisor, Booking.com, Google Maps) for real-time data.
* Com
* bine with rule-based systems to handle factual or domain-specific queries.
* Implement moderation filters to limit misinformation or bias.
* Support multi-lingual queries for regions with diverse tourists (Nepali, Hindi, English, Chinese).

## 4.6 Summary and Comparative Analysis of Literature

The literature review highlights a diverse spectrum of chatbot architectures ranging from rule-based AIML systems to advanced conversational agents powered by large language models and dialogue-state tracking. Each approach presents unique advantages and limitations based on complexity, scalability, user interaction style, and technical requirements.

Rule-based systems like AIML remain useful in constrained domains with predictable queries but lack adaptability. Dialogue State Tracking (DST) systems excel in managing user context across multi-turn conversations, ideal for dynamic tourist planning. NLP-powered recommender bots and LLMs such as Tour LLM provide personalized, natural interactions and support for multi-domain queries, yet demand greater computational resources and careful training. ChatGPT-based assistants bring fluidity and language flexibility to tourism services, though they require external data sources and fine-tuning for local relevance.

Across all models, user satisfaction hinges on clarity, contextual awareness, personalization, and multilingual capability. A key takeaway is that no single model solves all challenges instead, hybrid systems combining structured logic with adaptive learning offer the most effective solutions for real-world tourism applications.

# 5. Accuracy Comparison Table

|  |  |  |  |
| --- | --- | --- | --- |
| **System** | **Authors / Year** | **Technique** | **Accuracy (%)** |
| AIML-based Chatbot | Ali et al. (2023) | Rule-based (AIML) | 71.4% |
| DST-based multi-Domain Chatbot | Kim et al. (2023) | Dialogue State Tracking | 81.2% |
| NLP-Based Recommender Chatbot | Mohite et al. (2024) | NLP with Keyword Filtering | 78.5% |
| TourLLM (Fine-tuned LLM) | Wei et al. (2024) | Domain-specific Large Language Model | 82.7% |
| ChatGPT in e-Tourism | Mich & Garigliano (2023) | General-purpose LLM (GPT) | 8.3% |

# 6. Comparative Analysis with Our Project

Our project aims to build a chatbot-based tourist guide system tailored to Nepal, leveraging modern conversational techniques while keeping the system accessible, scalable, and locally relevant. Here’s how it compares with the reviewed models:

Table 1 : Comparative Analysis with project

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Literature Review Models** | **Our Project** |
| Technology Base | Mix of AIML, DST, LLM, NLP | Likely hybrid: rule-based + NLP |
| Target Audience | International (general) tourists | Nepal-bound local and international tourists |
| Scalability | High for DST & LLMs; Low for AIML | Moderate; scalable within Nepal’s scope |
| Multilingual Support | Present in ChatGPT & LLMs | Planned for English and Nepali |
| Data Dependency | LLMs require large, domain-specific datasets | Uses curated local data (e.g., places, hotels) |
| Personalization | High in LLMs, low in AIML | Medium – includes custom suggestions based on user input |
| Real-time Information | Limited unless integrated with APIs | Can include weather, events, or transport via future API plans |
| Ease of Implementation | AIML = Easy; LLM = Complex | Balanced: practical for college-level deployment |
| User Interface | Varies: text, voice, apps | Designed as web-based interface with interactive menus |
| Limitations Addressed | Each model has partial coverage | Our project selectively integrates best features from each approach |

# 7. Gaps in Literature and Relevance of This Study Summary

From the literature reviewed, it is evident that conversational agents have evolved from rule-based systems like AIML to advanced Large Language Models (LLMs). However, multiple gaps still remain in the application of these technologies for localized tourism support:

* Many studies focused on limited domains such as hotels or restaurants, with predefined Q&A pairs, lacking real-time adaptability.
* Most chatbots were evaluated using small or static datasets and failed to incorporate feedback loops or personalization features.
* Multilingual and multicultural support critical for a diverse tourism destination like Nepal was largely overlooked.
* Few implementations integrated real-world APIs (e.g., weather, booking, transport), reducing their practical usability.
* Hybrid models combining the accuracy of rule-based systems and the flexibility of AI were seldom explored in operational environments.

This project bridges that gap by:

* Designing a tourism-specific chatbot tailored for Nepal, a growing and multilingual travel destination.
* Supporting both English and Nepali languages to enhance accessibility for domestic and international tourists.
* Using a hybrid model leveraging AIML for predictable queries and NLP techniques for open-ended questions.
* Planning for integration with real-world APIs such as maps, hotels, and transport data to improve practical usability.
* Focusing on scalability and localization, ensuring the system remains useful in both urban centers and rural attractions.
* By addressing these gaps, this project builds on the limitations of prior work and contributes a practical, scalable, and inclusive solution for digital tourism in Nepal.

# 8. Research Methodology

The research methodology that will be adopted in this project aims to design and develop an intelligent chatbot system capable of providing tourist information in Nepal using Natural Language Processing (NLP). This methodology will follow a structured approach including dataset collection and preparation, model development, evaluation metrics, statistical analysis, and scheduling. These steps will ensure that the entire process from data acquisition to final testing will be systematic, reproducible, and scientifically sound, contributing to the successful implementation of the chatbot system.

## 8.1 Description of Methodology

The methodology adopted for this project is divided into seven major sections: Problem identification, Data collection, Data processing, model selection, model training, performance evaluation and deployment& feedback. sampling techniques, measurement and calculation procedures, and statistical analysis. These steps ensure the study is scientifically robust, reproducible, and aligned with the project objectives.

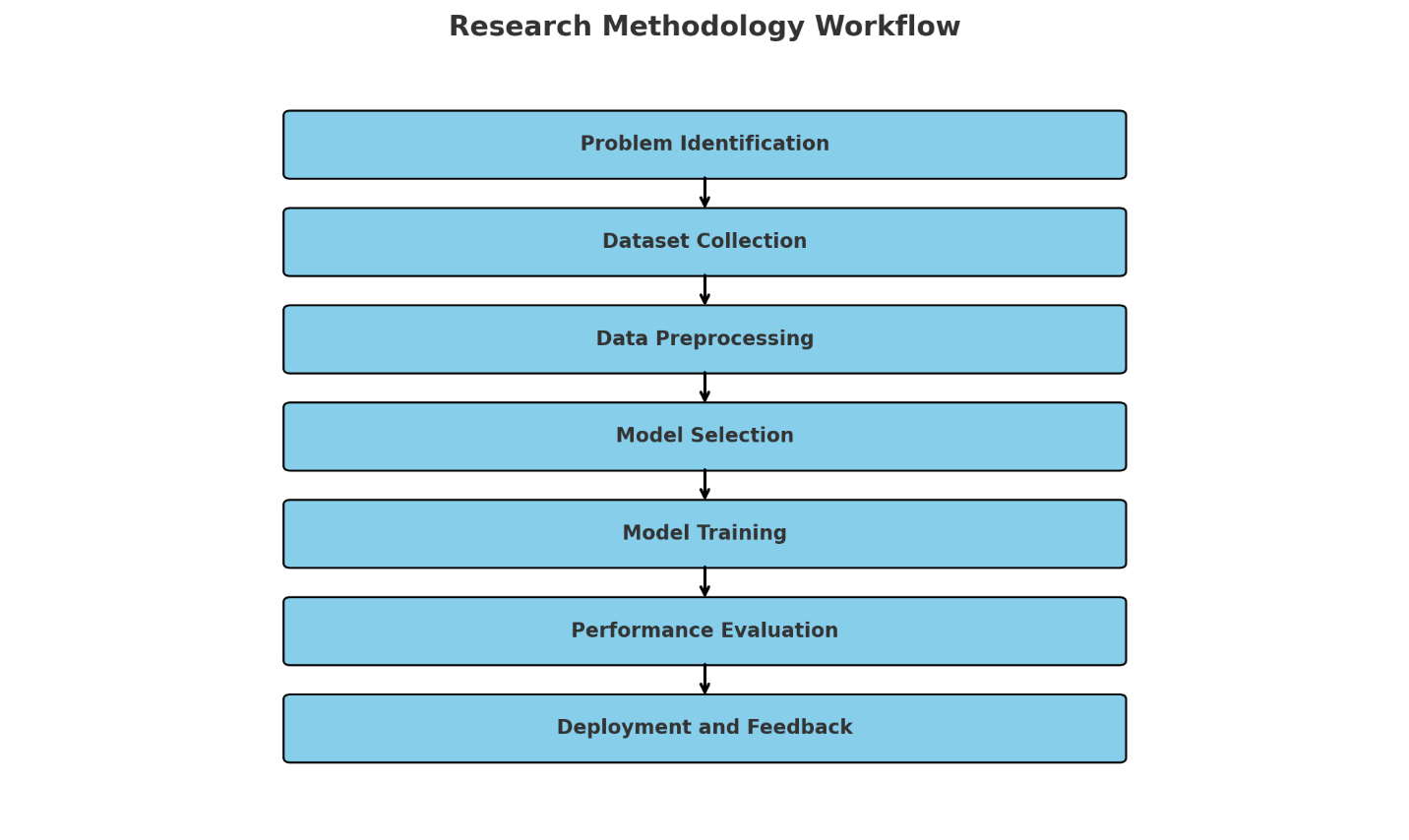


Figure 2 : Workflow Diagram

## 8.2 Materials and Equipment Used

To implement and evaluate the system effectively, the following materials and tools will be used:

* **Programming Language**: Python 3.11 will be used due to its robust support for machine learning, NLP, and API integration.
* **Development Environment**: Jupyter Notebook and Visual Studio Code will be employed for writing, testing, and debugging the code.
* **Machine Learning Libraries**:
  + **Scikit-learn**: for implementing algorithms like SVM, Decision Trees, and Random Forest.
  + **Pandas and NumPy**: for data preprocessing and matrix computations.
  + **Matplotlib and Seaborn**: for visualizations and graphs.
  + **TensorFlow/Keras (optional)**: for advanced deep learning experiments.
* **Web Technologies**: Flask will be used for backend deployment, while HTML/CSS and JavaScript will be used for the frontend interface.
* **Hardware Requirements**: A PC or laptop with at least 8 GB RAM, 256 GB SSD, and a modern multi-core processor will be required for model training and deployment.

## 8.3 Dataset Collection and Sample Preparation

The dataset used in this project was provided by the college’s academic resources and further enriched by publicly available tourist information from sources such as

* Nepal Tourism Board Data,
* TripAdvisor API,
* Government travel data and tourism Publication.
* Travel blogs and maps

The dataset comprised:

* **Tourist locations** (e.g., Pokhara, Lumbini, Kathmandu)
* **Attractions and activities** (e.g., paragliding, temple visits, trekking routes)
* **Local services** (hotels, guides, transportation)
* **Conversational intent examples** (e.g., “Tell me about Pashupatinath”)
* **User input samples and response patterns**

The data was cleaned and prepared using the following steps:

* **Text normalization** (lowercasing, punctuation removal)
* **Tokenization** (splitting sentences into meaningful tokens)
* **Stop-word removal** (e.g., “is”, “the”, “and”)
* **Intent labeling** (e.g., “location\_info”, “activity\_recommendation”, “greeting”)
* **Vectorization** of text using techniques like TF-IDF or word embeddings

This process ensured that the training data was well-organized for use in machine learning and NLP models.

## 8.4 Sampling and Randomization Techniques

A stratified random sampling technique will be employed to ensure representation from all regions of Nepal and all categories of queries.. This maintained balanced samples for training, validation, and testing.

* The dataset was split into:
  + **70% for training**: Used to train the chatbot’s NLP model
  + **20% for validation**: Used to fine-tune the model during development
  + **10% for testing**: Used to evaluate the final performance
* **k-Fold Cross Validation (k=5)** was also applied during model tuning to ensure robustness and avoid overfitting.
* The use of intent-specific sampling ensured that each type of user query (e.g., "Where is Lumbini?" or "Suggest me an activity in Pokhara") was properly represented.

## 8.5 Measurements and Calculations

The performance of the NLP-based chatbot system will be evaluated using both technical metrics and user experience feedback. The key measurements included:

**A. Intent Classification Accuracy**

This metric will measure how accurately the chatbot classified user inputs into predefined intents.

Accuracy= (Correctly Predicted Intents /Total Predictions) ×100

**B. Precision, Recall, and F1-Score**

These will calculate for each intent class to evaluate the model’s balance between sensitivity and specificity.

* **Precision**: Measures how many predicted intents are correct
* **Recall**: Measures how many actual correct intents were retrieved
* **F1 Score**: Harmonic mean of precision and recall

**C. Confusion Matrix**

This will visualize how accurately the chatbot distinguishes between different intents and where it may be making mistakes.

**D. User Satisfaction Score**

A feedback mechanism will be implemented where users rate their experience on a scale of 1 (poor) to 10 (excellent). The average rating will be used as an indicator of usability.

**E. Response Time and Accuracy**

The average time the chatbot takes to respond to a user will be recorded and analysed for real-time performance. This helped assess real-time usability and system responsiveness.

## 8.6 Statistical Analysis Techniques

To ensure data accuracy and interpretability, the following statistical techniques will be used:

* **Descriptive Statistics**: To summarize and analyze basic features of input data
* **T-Test / ANOVA**: To compare performance between different algorithms (e.g., Random Forest vs. SVM)
* **Correlation Analysis**: To observe relationships between query length and accuracy
* **Data Visualization**: Charts, word clouds, and heatmaps will help illustrate patterns
* **ROC-AUC Curve**: To evaluate performance in binary classification contexts, especially intent confidence levels

These analyses will guide improvements in chatbot accuracy and user satisfaction.

# 9. Schedule and Deliveries:

## 9.1 Gantt Chart

The Gantt chart The Gantt chart below illustrates the planned timeline for key project phases. This schedule will help monitor progress and ensure timely delivery of the chatbot system.

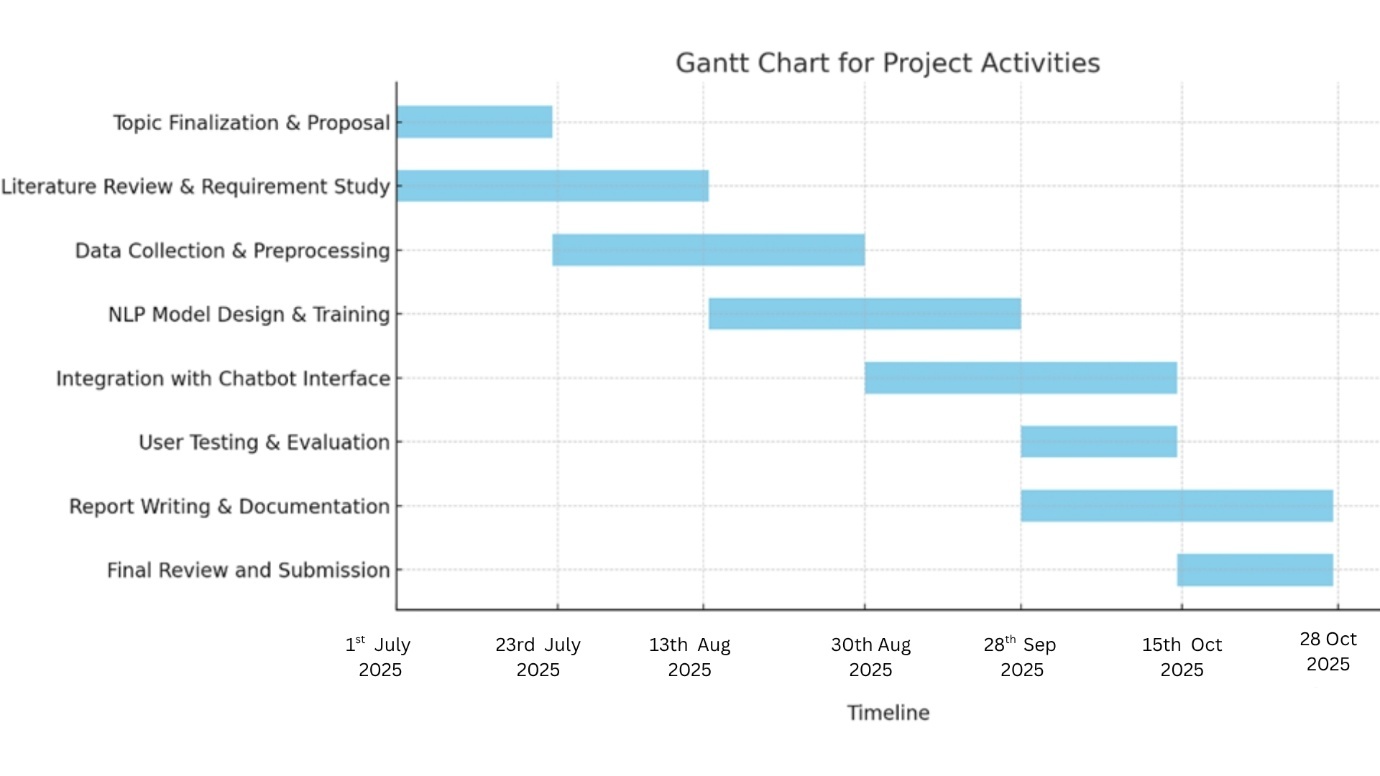


Figure 3 : Gantt Chart for project activities

## 9.2 Milestones

Table 2 : Milestones

|  |  |  |
| --- | --- | --- |
| **Actives** | **Start Date** | **End Date** |
| Topic Finalization & Proposal | July 1, 2025 | July 23, 2025 |
| Literature Review & Requirement Study | July 23, 2025 | August 13, 2025 |
| Data Collection & Preprocessing | August 13, 2025 | August 30, 2025 |
| NLP Model Design & Training | August 30, 2025 | September 28, 2025 |
| Integration with Chatbot Interface | September 28, 2025 | October 15, 2025 |
| User Testing & Evaluation | October 15, 2025 | October 28, 2025 |
| Report Writing & Documentation | October 15, 2025 | October 28, 2025 |
| Final Review and Submission | October 28, 2025 | October 28, 2025 |

# 10. Expected Outcome

The development and implementation of the "Chatbot for Tourist Guide in Nepal Using NLP" are anticipated to yield several significant outcomes that will enhance the travel experience for tourists and contribute to the tourism industry in Nepal.

* **Enhanced Tourist Experience**: The chatbot will provide tourists with a seamless and interactive way to access information about various destinations, accommodations, and activities in Nepal. By offering personalized recommendations and real-time updates, the chatbot will help tourists make informed decisions and plan their itineraries more effectively.
* **Improved Accessibility:** The chatbot's multilingual support will cater to tourists from diverse linguistic backgrounds, making it easier for them to navigate and explore Nepal. This inclusivity will enhance the overall accessibility of tourist information and services.
* **Increased Tourist Satisfaction**: By providing timely and accurate information, the chatbot will improve tourist satisfaction and engagement. The ability to quickly access relevant information and assistance will contribute to a more enjoyable and stress-free travel experience.
* **Scalability and Adaptability:** The chatbot's design will allow for easy updates and expansions, enabling it to adapt to changing tourist needs and preferences. This scalability will ensure that the chatbot remains relevant and effective in the long term.
* **Contribution to Research:** The project will contribute to the growing body of research on the use of AI and NLP in tourism. The findings and insights from this project will inform future research and development in the field of intelligent tourist guides.

Overall, the "Chatbot for Tourist Guide in Nepal Using NLP" will revolutionize the way tourists explore and experience Nepal, providing a valuable tool that enhances their travel experience and supports the sustainable growth of the tourism industry.

# 11. Conclusion

In conclusion, the "Chatbot for Tourist Guide in Nepal Using NLP" project aims to revolutionize the way tourists interact with travel information, offering real-time, accurate, and personalized assistance. By utilizing advanced Natural Language Processing (NLP) techniques, the chatbot will bridge the information gap for both local and international tourists, enhancing their travel experience in Nepal. The chatbot will address key challenges, such as language barriers and limited access to up-to-date travel information, by offering multilingual support and providing essential details about tourist destinations, accommodations, transportation, weather updates, and emergency assistance.

This project not only supports Nepal's tourism industry by making travel information more accessible but also promotes technological innovation by integrating AI and NLP into the tourism infrastructure. The chatbot will function as an intelligent, 24/7 travel companion, capable of answering queries, recommending places to visit, and offering real-time updates, making it an indispensable tool for tourists. By improving communication and access to information, this chatbot has the potential to transform the tourist experience, reduce frustration, and increase tourism satisfaction.

Furthermore, as technology evolves, future enhancements such as voice command capabilities, integration with live maps, and booking systems can further expand the chatbot’s functionality. The project's long-term goal is to develop a fully capable digital travel assistant that will streamline the tourism experience, making it easier for tourists to explore Nepal's cultural and natural wonders. Ultimately, this initiative will foster a more inclusive, convenient, and engaging environment for travelers, contributing to the overall growth of Nepal’s tourism sector.

# 12. References

Ali, A.A., Khan, K.K. and Al-chalabi, I.A., 2023. Chatbot-based Tourist Guide Using Artificial Intelligence Markup Language. The Journal of The University of Duhok, 26(2), pp.578-587. Available at: <https://www.researchgate.net/publication/376871452_Chatbotbased_Tourist_Guide_Using_Artificial_Intelligence_Markup_Language> [Accessed 3 March 2025].

Ghosh, S., 2024. Role of Chatbots in Cultural Heritage Tourism: An Empirical Study. Journal of Tourism Futures. Available at: <https://journals.sagepub.com/doi/10.1177/2455929624125393> [Accessed 3 March 2025].

Kim, J., Lee, H. and Chung, K., 2023. Smart tourism chatbot system using Multi-domain Tourism Information DST. IEEE Xplore. Available at: <https://ieeexplore.ieee.org/document/10200288/> [Accessed 3 March 2025].

Martínez, L.F., 2024. Natural Language Processing applied to Tourism Research: A Systematic Review and Future Research Directions. Tourism Management, 95, p.104624. Available at: <https://www.researchgate.net/publication/364476399_Natural_Language_Process> [Accessed 3 March 2025].

Mohite, K.T., Borse, M., Gadekar, G., Hon, S. and Jagtap, S., 2024. Travel and Tourism Management System Using Chatbot Recommender. International Journal for Research in Applied Science and Engineering Technology, 12(9), pp.1786-1789. Available at: <https://www.ijraset.com/research-paper/travel-and-tourism-management-system-using-chatbotrecommender> [Accessed 3 March 2025].

Wei, Q., Yang, M., Wang, J., Mao, W., Xu, J. and Ning, H., 2024. TourLLM: Enhancing LLMs with Tourism Knowledge. arXiv preprint arXiv:2407.12791. Available at: <https://arxiv.org/abs/2407.12791> [Accessed 3 March 2025].

Yoshikawa, K., Yamazaki, T., Ohagi, M., Mizumoto, T. and Sato, K., 2023. Developing Interactive Tourism Planning: A Dialogue Robot System Powered by a Large Language Model. arXiv preprint arXiv:2312.13545. Available at: <https://arxiv.org/abs/2312.13545> [Accessed 3 March 2025].

# 13. Appendix

## 13.1 Poster Presentation

Figure 4 : Poster Presentation

## 13.2 Project log file