SUMMER INTERNSHIP REPORT

Computer Science and Engineering

JIIT Noida

B TECH – Ist Year

(2-Credit Internship)

Travel Buddy: Travel Recommendation System

Submitted by

Karvy Singh	2401030234
Tanishka Singh	J992401210097
Harshit Shukla	2401200047
Jayant Singh	J992401040089
Prasoon Sharma	J992401210051

Internship Duration: June 9 - July 5, 2025

Submitted to

Dr. Neetu Singh

Dr. Imraan Rasheed

Dr. Kirti Aggarwal

Declaration

I hereby declare that the Summer Internship Report titled "Travel Buddy: Travel Recommendation System" is the outcome of my own work, carried out under the guidance of-
Dr. Neetu Singh
Dr. Imraan Rasheed
Dr. Kirti Aggarwal
and has not been submitted elsewhere for any other academic purpose.

Name of Student: Karvy Singh

Signature: _____

Date: _____

Certificate

This is to certify that Karvy Singh, Roll No. 2401030234, has successfully completed the Summer Internship Project titled "Travel Buddy: Travel Recommendation System" during the period June 9 – July 5, 2025 under my supervision.

Mentor's Name: _____

Designation: _____

Department: _____

Organisation: _____

Signature: _____

Acknowledgement

I express my sincere gratitude to my mentor-

Dr. Neetu Singh

Dr. Imraan Rasheed

Dr. Kirti Aggarwal

for their invaluable guidance and support throughout this internship. I also thank Jaypee Institute of Information Technology for the opportunity to work on this project, and my family and friends for their encouragement.

Contents

0.1   Introduction

Travelers often find it overwhelming to search for suitable tourist spots due to the abundance of options available. The diversity in preferences and lack of personalized suggestions further complicates their decision-making process, so for more personalized recommendations, so we made recommendation system that uses IR+ML to give you details of travel destinations fromthe our dataset consisting of 3000 entries, along with some pictures from net.

0.2   Objectives of the Internship

- Understand Core IR Concepts

  - Study and implement ranking and retrieval metrics.

  - Analyze search engine behavior and document relevance.

- Apply Machine Learning to IR Tasks

  - Use supervised models to improve ranking or classification.

  - Train and evaluate models on real-world datasets.

- Explore Feature Engineering and Model Evaluation

  - Extract meaningful features (e.g., TF-IDF).

  - Evaluate models with appropriate metrics (e.g., MAP, NDCG, F1).

- Build End-to-End IR or ML Pipelines

  - Preprocessing, model training, evaluation, and reporting.

  - Automate workflows where possible (e.g., using scikit-learn or PyTorch).

- Gain Practical Experience

  - Work with real datasets (text, tabular, etc.).

  - Contribute to team projects or research-oriented tasks.

0.3   Company/Problem Overview

Sites like MakeMyTrip focus on hotels, prices, journey but we lack sites recommending which place to visit as per your interest, for that we need to use old-school google search, which does not allow you to put in filters for more personalized recommendations, so we made recommendation system that uses IR+ML to give you details of travel destinations from the our dataset, with photos from the net, along with simple sober and yet better interface for interaction.

0.4   Tools and Technologies Used

- Neovim editor

- Python

  - flask==2.3.3

  - numpy==1.26.4

  - pandas==2.2.2

  - scikit-learn==1.4.2

  - scipy==1.13.1

  - xgboost==2.0.3

  - requests==2.32.3

- Deployment app

0.5 Weekly Progress Summary

- Week 1: Foundations of Information Retrieval Studied IR systems, text preprocessing (tokenization, stopword removal, stemming, lemmatization), inverted index construction, retrieval models (Boolean, Vector Space, TF-IDF), similarity metrics (Cosine, Jaccard, Euclidean), and evaluation metrics (Precision, Recall, F1-score, MAP, NDCG).

- Week 2: Machine Learning for IR Explored the role of ML in IR (ranking, classification, recommendations), implemented Naive Bayes for document classification, used Scikit-learn for model building, and studied feature engineering (lexical, semantic, behavioral).

- Week 3: Learning to Rank and Recommendation Systems Learned about learning-to-rank approaches (pointwise, pairwise, listwise), RankNet and LambdaMART, embedding-based retrieval (word/sentence embeddings), and recommendation techniques (content-based, collaborative filtering).

  Week 4: Project Week Built an end-to-end IR+ML system including text preprocessing, TF-IDF ranking, document classification, and evaluation using IR metrics. Integrated concepts from previous weeks into a working prototype.

0.6 System Design / Workflow / Architecture

0.7 Implementation Details

0.8   Screenshots / Code Snapshots / Output Samples

0.9 Challenges Faced and Solutions

0.10 Key Learnings

1. **Foundations of Information Retrieval**

   - Understood how IR systems function, focusing on preprocessing techniques such as tokenization, stopword removal, stemming, and lemmatization.

   - Built an inverted index for efficient document retrieval.

   - Applied retrieval models including Boolean, Vector Space, and TF-IDF.

   - Used similarity metrics like Cosine, Jaccard, and Euclidean distance to assess document relevance.

   - Evaluated system performance using metrics such as Precision, Recall, F1-score, MAP, and NDCG.

2. **Machine Learning for IR**

   - Explored how ML enhances IR through classification and ranking tasks.

   - Implemented Naive Bayes for document classification using Scikit-learn.

   - Conducted feature engineering using lexical, semantic, and behavioral features to improve model effectiveness.

3. **Learning to Rank and Recommendations**

   - Studied learning-to-rank methods: pointwise, pairwise, and listwise approaches.

   - Learned ranking models such as RankNet and LambdaMART.

   - Applied embedding-based retrieval using word and sentence embeddings for semantic search.

   - Developed recommendation systems using content-based and collaborative filtering techniques.

4. **End-to-End System Development**

   - Designed and implemented a full IR+ML pipeline, including preprocessing, ranking, classification, and evaluation.

   - Integrated diverse models into a functional travel recommendation application.

   - Focused on user experience by building a clean, intuitive interface for personalized destination suggestions.

0.11 Conclusion

Over four weeks, we explored the core principles and practical applications of Information Retrieval (IR) and Machine Learning (ML) to build a travel recommendation system that goes beyond conventional platforms like MakeMyTrip. Unlike traditional travel sites that focus mainly on logistics, our system focuses on personalized destination recommendations based on user interests, using techniques like TF-IDF ranking, classification models, and similarity metrics. By integrating IR techniques with ML-driven classification and ranking algorithms, we created a user-friendly prototype that retrieves relevant travel destinations and displays them with images and contextual insights, enhancing user experience and engagement.

0.12   References

Bibliography

[1] Dataset. https://www.kaggle.com/datasets/naqibahmedkadri/famous-indian-tourist-places

[2] Flask Documentation. https://flask.palletsprojects.com/

[3] XGBoost Documentation. https://xgboost.readthedocs.io/

[4] NumPy Documentation. https://numpy.org/doc/

[5] Pandas Documentation. https://pandas.pydata.org/docs/

[6] SciPy Documentation. https://docs.scipy.org/doc/scipy/

[7] Scikit-learn Documentation. https://scikit-learn.org/stable/documentation.html

0.13   Appendix