A Major Project Final Report on

Discovering Nepal

Submitted in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Engineering in Information Technology

Under Pokhara University

Submitted by:

Aayush Lamsal, 181503

Aashutosh Bista, 181501

Rabin Gurung, 181426

Under the supervision of

Associate Professor Mrs. Resha Deo

Date:

2023/07/27

Department of Information Technology

NEPAL COLLEGE OF

INFORMATION TECHNOLOGY

Balkumari, Lalitpur, Nepal

ACKNOWLEDGEMENT

First, we would like to express our sincere gratitude towards the Department of IT Engineering, Pokhara University for including the 8th semester major project as a part of the syllabus.

After a rigorous and continuous united effort, we are glad to present this report as an outcome of the major project report on "Discovering Nepal". We express our sincere gratitude to those helping hands that come forward and direct us towards the success of the whole project with the desired output. We extend our genuine thankfulness to our project supervisor Mrs. Resha Deo for her valuable and great support and guidance in all the happenings regarding the project. We must express our thanks to **Dr. Roshan Chitrakar,** project coordinator for his encouragement and guidance throughout the project. We are indebted to the department of information technology for providing support to add on our venture. Last, but not the least, we would like to thank our teachers and colleagues who have been knowingly or unknowingly part of this project and lent support and views during the entire development time.
Every attempt has been made to include each and every expectation of the project in this report so that readers can clearly understand the project. We would be pleased to get feedback on it.

Sincerely,
Aayush Lamsal(181503)
Aashutosh Bista (181501)
Rabin Gurung (181426)

ABSTRACT

Tourism is a significant contributor to the economic growth of many countries, including Nepal. However, many tourist sites in Nepal go unnoticed, leading to an uneven distribution of economic benefits from tourism. This project, Discovering Nepal, aims to develop a recommendation system using content-based filtering, top-Rated recommendations and nearby recommendations of tourist sites to visitors. The recommendation system will provide personalized recommendations based on the user's location and preferences.

By using machine learning algorithms, the system can help visitors discover new places to visit, increase awareness of lesser-known tourist sites in Nepal, and promote their visitation. The project has the potential to contribute to economic growth and development of more tourism centers in the country. The literature study for this project will include research on machine learning algorithms for recommendation systems and previous studies on similar projects done such as Holidify and Tripadvisor to some extent. The proposed methodology involves Agile Methodology with steps including data collection and preparation, data preprocessing, feature extraction, recommendation engine development and web development. The tools and systems involved in development of this project involve Jupyter Notebook, HTML, CSS, Javascript, Ajax, PHP, Python and libraries. The limitations of the project include the availability and quality of data on tourist sites in the area and the accuracy of recommendations[1]. Overall, the project aims to provide a solution to the problem of an uneven distribution of economic benefits from tourism in Nepal by utilizing machine learning algorithms to recommend personalized tourist sites.

Keywords: *Recommendation System, Content Based Filtering, Machine Learning Algorithm, Agile Methodology, Web Development*.

Table Of Contents:

INTRODUCTION

Tourism is an important industry in Nepal and is a major contributor to its economy. Nepal has a diverse range of tourist attractions, including "Religious and Pilgrimage sites", "Cultural and Historical sites", "Lakes, Rivers and Water Bodies", etc. However, despite the abundance of tourist sites, many of these sites go unnoticed, while only a few popular sites receive the majority of visitors. This uneven distribution of visitors to tourist sites can lead to economic disparities between regions, where only popular sites benefit from tourism, while lesser-known sites remain underdeveloped and unexplored.

To address this issue, this project aims to develop a web application, Discovering Nepal, which uses machine learning algorithms to recommend even not so popular sites while recommending places to visit on the basis of user preference. The application will utilize content-based filtering, K-Nearest Neighbor algorithm to provide personalized recommendations based on the user's location, preferences, and other factors such as top rated places. The application will be designed to help visitors discover new places to visit and increase awareness of lesser-known tourist sites in Nepal, promoting their visitation.

Visitors can create their own profiles and provide information such as their location, interests, and previous visits to tourist sites. This information will be used to personalize recommendations and ensure visitors receive suggestions tailored to their preferences.

The web application will also be beneficial for local tourism agencies and businesses that can use the recommendation system to promote lesser-known tourist sites in their area, leading to economic growth and development of more tourism centers in the country through suggestion boxes for promoting new places. Furthermore, the data collected from the web application can be used to identify trends in visitor preferences and help tourism businesses make data-driven decisions to improve the visitor experience.

.

PROBLEM STATEMENT

Tourism is a vital industry in Nepal and a significant contributor to its economy. Despite having a diverse range of tourist attractions, many tourist sites go unnoticed, while only popular and trending sites receive the majority of visitors. This uneven distribution of visitors to tourist sites can lead to economic disparities between regions, where only popular sites benefit from tourism, while lesser-known sites remain underdeveloped and unexplored.

The lack of awareness of these lesser-known tourist sites is due to several reasons, including a lack of proper information and guidance on where to go, what to see, and how to get there. This problem is compounded by the lack of a centralized system for collecting and sharing information on tourist sites in Nepal. The existing system that provides information on tourist sites in Nepal is mainly social media. There isn't a website that is specifically designed for providing personalized recommendations of places in Nepal.

To address this problem, this project aims to develop a web application, Discovering Nepal, that uses machine learning algorithms to recommend tourist sites to visitors under categories top-rated places, nearby places and recommended places. The application will utilize content based filtering and KNN algorithm to provide personalized recommendations based on the user's location and preferences. The application will be designed to help visitors discover new places to visit and increase awareness of lesser-known tourist sites in Nepal, promoting their visitation.

PROJECT OBJECTIVE

1. Develop a recommendation system that suggests tourist sites based on user preferences and location.
2. Help promote the local businesses as well as not so popular tourist spots.
3. Ask for user suggestions in order to promote more places.

SIGNIFICANCE OF STUDY

1. Increase the economic benefits of tourism by promoting the visitation of lesser-known tourist sites in Nepal.
2. Encourage the development of more tourism centers in the country by increasing the awareness and popularity of these sites.
3. Provide visitors with personalized recommendations that improve their experience and satisfaction with Nepali tourism.
4. Contribute to the development of machine learning and recommendation system techniques in the field of tourism.

SCOPE AND LIMITATIONS

The scopes of **Discovering Nepal** which are stated as below:

1. The web application will provide recommendations for tourist sites in Nepal.
2. The recommendation system will use content based filtering to personalize recommendations for individual users.
3. The application will provide a platform for users to suggest admin with new spots to promote.
4. While recommending the places to visit on the basis of user preference, the system will also recommend nearby places to visit that may or may not be that popular with the aim of promoting not so popular places as well.
5. This project can help in solving the problem of uneven distribution of tourists in Nepal resulting in economic growth of local businesses and the whole country.

The limitations of **Discovering Nepal** are as follows:

1. The web application will require internet access to function, which may limit its accessibility in some areas of Nepal.
2. The recommendation system may not be able to provide accurate recommendations for users with unique preferences or requirements.
3. The collection of data on tourist sites may be limited by the availability of information and the resources available for data collection.
4. At the initial stage, due to lack of enough user information, recommendations may be suitable.

LITERATURE REVIEW

Several web-based applications have been developed to help travelers find the best tourist spots.These applications employ different approaches to provide recommendations to travelers.Some of the applications are Tripadvisor, Expedia Travel, Lonely Planet, etc.

One approach is collaborative filtering, where the system recommends tourist spots based on the ratings given by other users with similar preferences.A study by Lee et al.(2014) proposed a collaborative filtering algorithm that uses a hybrid method of clustering and matrix factorization to provide recommendations to travelers. Some applications which provide recommendations by using collaborative filtering are Google Maps, Tripadvisor, Expedia Travel.

Another approach is content-based filtering, where the system recommends tourist system recommends tourist spots based on the similarity of their attributes to those of the tourist spots previously visited by the user.A study by li et al (2016) proposed a content-based filtering approach that uses both the textual and visual content of the tourist spots to provide recommendations to travelers. System which provides recommendations by using content-based filtering is Lonely Planet.

A third approach is hybrid filtering,which combines collaborative filtering and content-based filtering to provide more accurate recommendations.A study by Yang et al (2020) proposed a hybrid filtering algorithm that uses a deep neural network to learn the features of tourist spots and provide recommendations to travelers.

PROPOSED METHODOLOGY

The proposed methodology for our project is **Agile Methodology.**

Fig: Agile Methodology

The basic explanation of steps involved in **Agile Methodology** are:

1. **Plan**: In the planning phase, the project team and stakeholders work together to identify the overall goals and objectives of the project. The team then creates a prioritized list of features and tasks, which is known as the product backlog. The product backlog is continually refined and updated throughout the project to reflect changes in priorities, feedback from stakeholders, and new ideas.
2. **Design**: During the design phase, the team plans a sprint (a fixed period of time, usually 2-4 weeks) and selects the tasks that will be completed during the sprint. The team estimates the time required for each task and assigns tasks to team members. The goal of this phase is to create a plan that is achievable within the timeframe of the sprint.
3. **Develop**: During the development phase, the team completes the selected tasks using iterative development and collaboration. The team works in short cycles, completing a small set of tasks and then reviewing and testing the work before moving on to the next

set of tasks. This approach allows the team to make adjustments and improvements as they go, based on feedback and new information.

4. **Test**: Testing is a critical part of the Agile methodology. The team tests the completed work to ensure that it meets the requirements and quality standards. The team may use automated testing tools and continuous integration to streamline the testing process and catch errors early.
5. **Deploy**: Once the work is completed and tested, it is deployed to a production or staging environment. This allows stakeholders to see the work in action and provide feedback.
6. **Review**: After the sprint is completed, the team and stakeholders review the work that was completed and gather feedback. This feedback is used to refine the product backlog and plan the next sprint. The review process is an important part of Agile methodology, as it allows the team to continuously improve and adjust their approach based on real-world feedback [].

SOFTWARE USED

1. HTML, CSS, Javascript
2. PHP, Ajax and MySQL
3. VSCode, Jupyter Notebook

MACHINE LEARNING:

Machine learning (ML) is a subset of Artificial Intelligence (AI) that focuses on developing algorithms and models that enable computers to learn from data, identify patterns, and make predictions or decisions without being explicitly programmed. In essence, it involves training a computer system to recognize patterns and make decisions based on those patterns, rather than relying on a predetermined set of rules.

Machine learning is important because it allows computers to learn from data and improve their performance over time, without the need for human intervention. This has many applications in fields such as image and Speech Recognition, Natural Language Processing (NLP), Fraud Detection, Personalized Recommendations, Autonomous Vehicles, etc.

There are several types of machine learning, including:

1. Supervised Learning: This involves training a model on labeled data, where the desired output is already known, in order to predict outcomes for new, unseen data. Examples include classification (predicting categorical labels) and regression (predicting continuous values).
2. Unsupervised Learning: This involves training a model on unlabeled data, where the desired output is unknown, in order to identify patterns or relationships within the data. Examples include clustering (grouping similar data points together) and dimensionality reduction (reducing the number of features or variables in the data).
3. Reinforcement Learning: This involves training a model to make decisions based on rewards or penalties received from its environment, in order to maximize its long term performance. Examples include game-playing agents and autonomous robots.
4. Semi-Supervised Learning: This involves training a model on a combination of labeled and unlabeled data, in order to improve its performance on the labeled data. This is often used when labeling large amounts of data is time-consuming or expensive.
5. Transfer Learning: This involves reusing pre-trained models or knowledge from one task to improve performance on another related task. This can save time and resources, and is often used in computer vision and natural language processing tasks.

Recommendation System

A recommendation system is a type of machine learning application that provides personalized recommendations to users based on their interests, preferences, and behavior. These systems are used in a variety of applications, such as e-commerce, social media, entertainment, and content streaming platforms.

The importance of recommendation systems lies in their ability to enhance user experience, increase engagement, and drive revenue for businesses. By analyzing user data, such as past purchases, search history, and user interactions, recommendation systems can suggest relevant products, services, or content that users are likely to enjoy, thereby increasing customer satisfaction and loyalty.

There are several types of recommendation systems, including:

1. Content-Based Filtering: This type of recommendation system suggests items that are similar to those that a user has liked in the past. It analyzes the characteristics of items, such as their metadata or textual descriptions, to identify patterns and make recommendations.
2. Collaborative Filtering: This type of recommendation system suggests items based on the preferences of similar users. It analyzes the behavior and preferences of multiple users to identify similarities and make recommendations.
3. Hybrid Recommendation Systems: These systems combine multiple approaches, such as content-based and collaborative filtering, to provide more accurate and personalized recommendations.

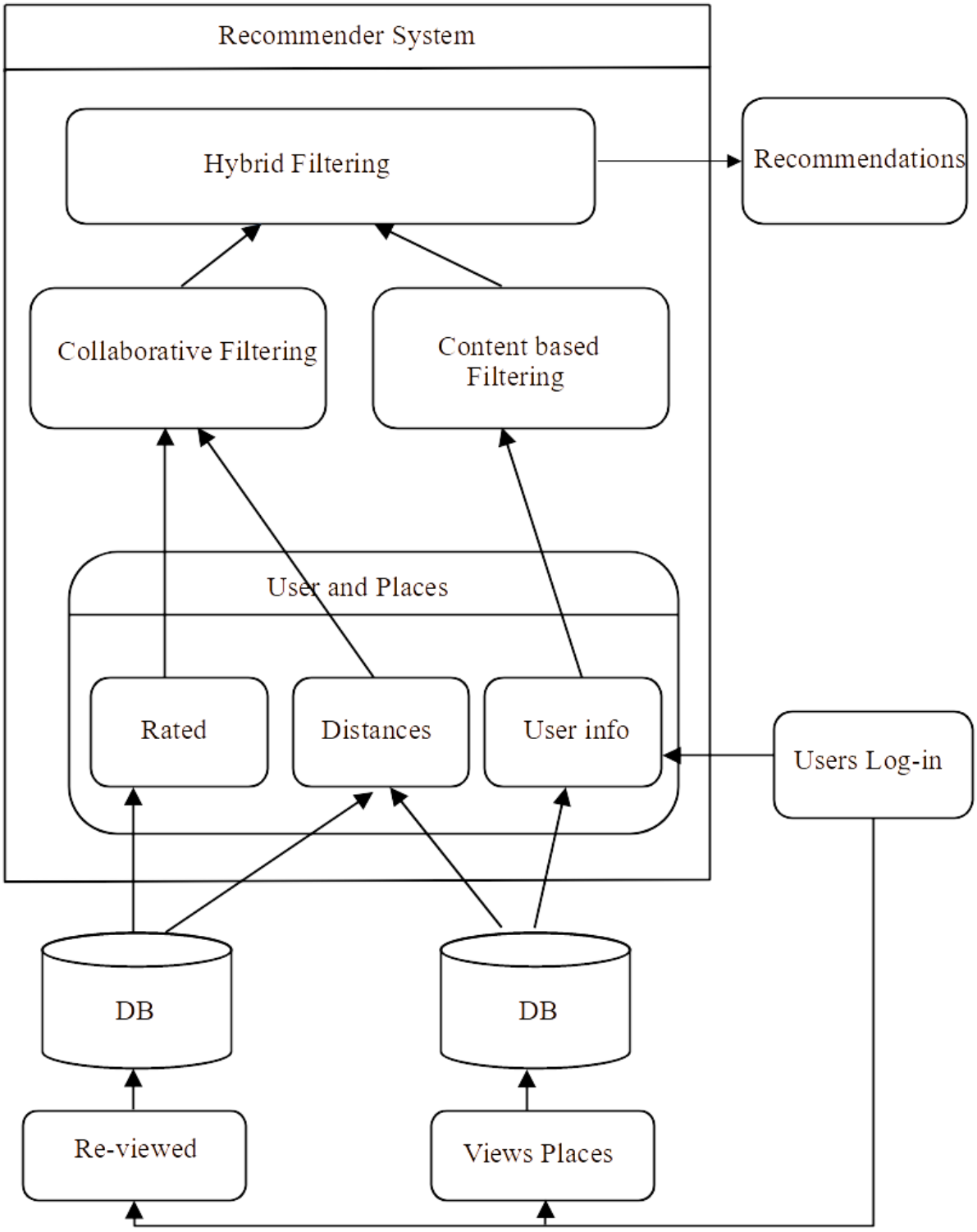


Fig: Discovering Nepal Recommendation System

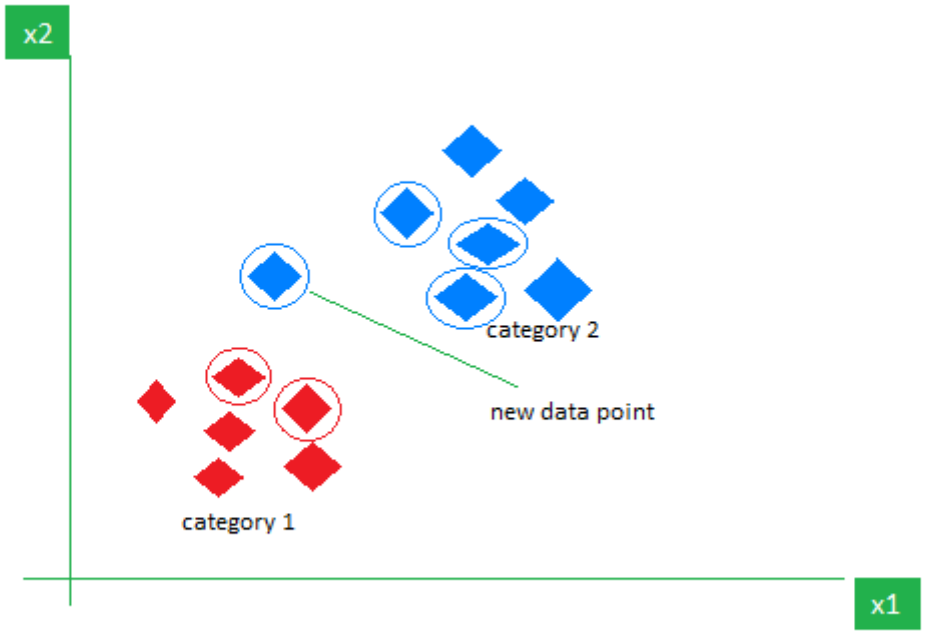
K-Nearest Neighbor Algorithm

In the context of recommendation systems, KNN can be used as both content based and collaborative filtering method. Content-based filtering is a recommendation technique used in recommender systems to make personalized recommendations to users based on the characteristics or content of items and the user's preferences. Instead of relying on the behavior and preferences of other users (collaborative filtering), content-based filtering focuses on the features or attributes of the items themselves. KNN works by identifying the K most similar users to a given user and recommending places that those similar users have liked or visited or wished to visit.

K-Nearest Neighbors (KNN) algorithm could be adapted for content-based filtering in a recommendation system:

The key steps would be:

1. Represent each place as a vector based on extracted keywords or attributes from the place description, reviews, etc.
2. When a user provides a query or context, convert that into a vector using the same process.
3. Find K most similar place vectors to the user's context vector based on distance metrics like cosine similarity.
4. The closest matches become the nearest "neighbors" in the vector space.
5. Recommend the top K places to the user.



Working in our system:

1) Data Collection: Gather data about the tourist spots, including attributes such as Category, Description, Major Attractions, Avg. Ratings, Total Ratings, Latitude, Longitude, and Popularity.
2) User Preference Acquisition: Obtain user preferences for the categories of tourist spots they are interested in. The user can specify up to three preferred categories.
3) Profile Creation: Create a user profile based on the user's preferred categories. The user profile represents the user's preferences for specific categories of tourist spots.
4) Attribute Extraction: Convert the tourist spots and the user profile into feature vectors by extracting relevant attributes such as Category, Description, and Major Attractions.
5) Similarity Calculation: Calculate the similarity between the feature vectors of the tourist spots and the user profile. This can be done using various similarity metrics like cosine similarity or Euclidean distance.

6) Top-Rated Recommendations: Recommend the user with top-rated tourist spots from their preferred categories. Consider both Avg. Ratings and Total Ratings count to ensure valuable recommendations.

7) Content-Based Filtering: When the user views any recommended tourist spot, use content-based filtering to recommend other similar tourist spots. Calculate the cosine similarity between the descriptions and major attractions of the viewed spot and other spots to identify similar ones.
8) Nearby Places Recommendations: If the user searches for a popular place, use the distance-based algorithm to recommend nearby places with moderate or unpopular popularity values to promote these less explored locations.

Content-based filtering is especially beneficial when users have unique preferences or when there is limited user data available. By analyzing the content and attributes of tourist spots, the algorithm can provide personalized and relevant recommendations to users, enhancing their overall tourism experience.

RESULT AND DISCUSSION

The followings results were obtained during the project development:

- The admin can update or add new places and destinations.
- The users can comment or review particular destinations after logging in.
- Based on users preferences, the destinations can be recommended.
- Users can make suggestions of new places to the system.

Further Works

The following tasks can be done to improve system:

- Collaborative filtering.
- Providing users with the ability to ask for description modification.

Use-Case Diagram:

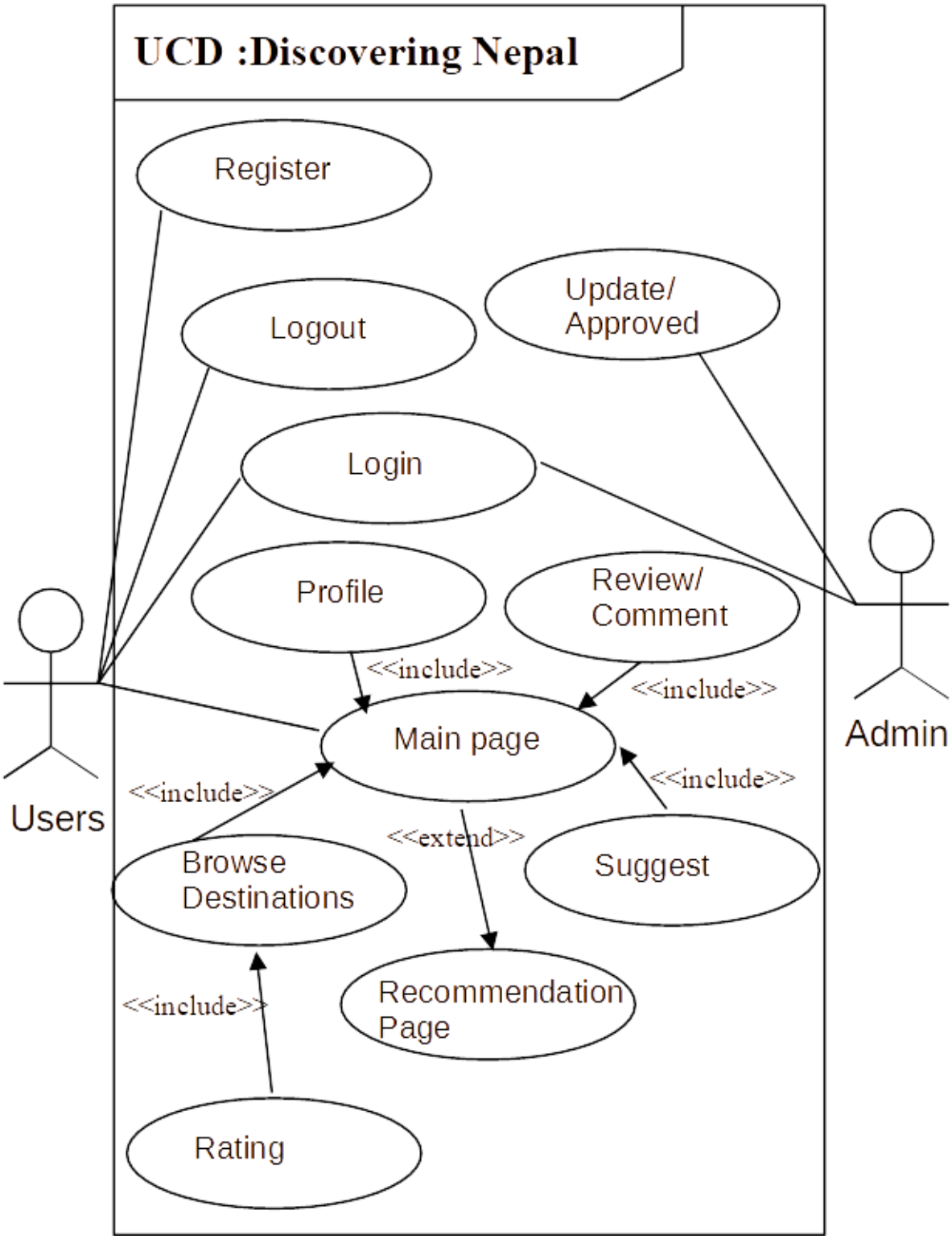


Fig: Use case Diagram of Discovering Nepal

System Sequence Diagram

Login SSD

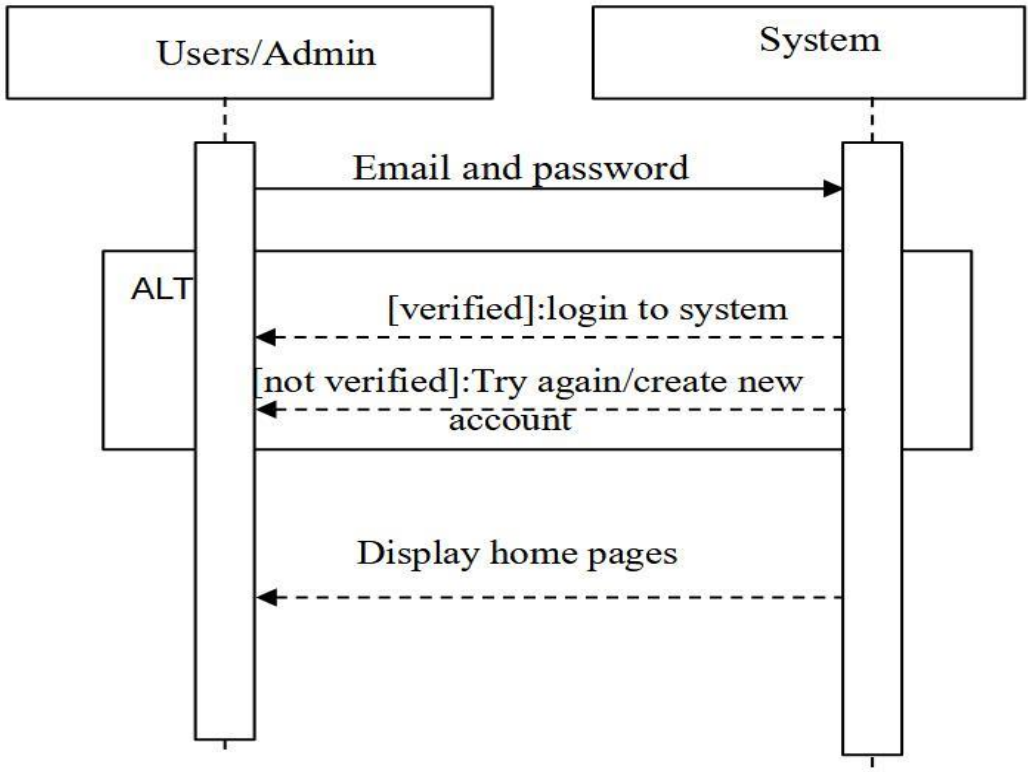


Fig: Login

Name: Login

Actor:Users, Admin

Description: The User and admin can login to the system before they use any of the functionalities.

Precondition:Users/Admin is identified and authenticated by the system

Postcondition: Userslogins to the system. Customers view places. Admin has functionalities like managing customers and merchants & viewing queries

Search SDD

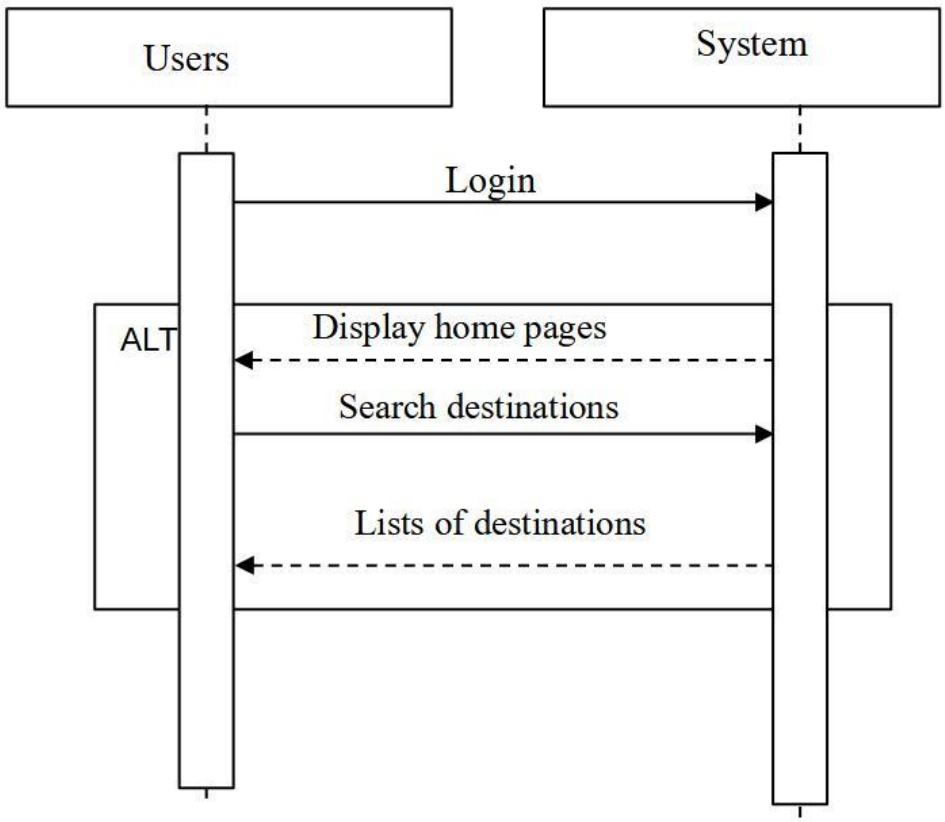


Fig: Search

Name: Search

Actor: Customer

Description: The User can search the places that he/she was interested in.

Precondition: The users search request is processed by the system

Postcondition: Users gets the search results. Customers view places.

Update

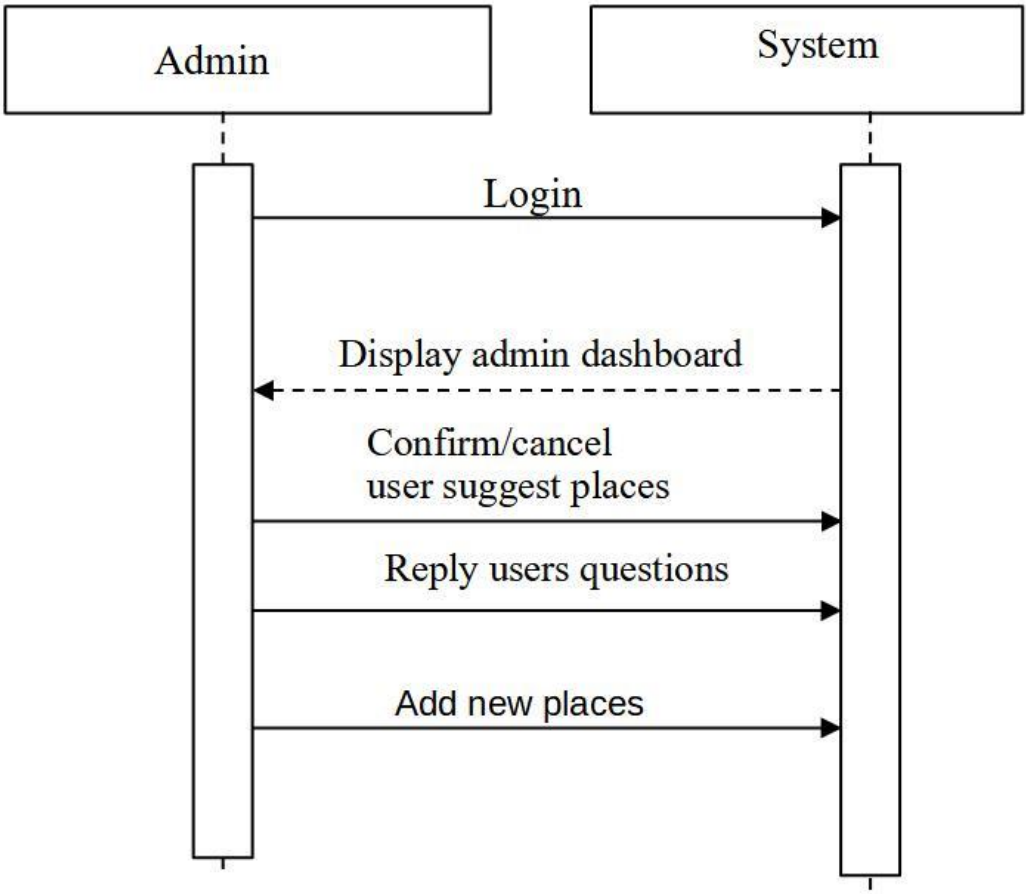


Fig: Update

Name: Update

Actor: Admin

Description: The admin can update/add new places to the system. The admin can also view the messages from users on the admin dashboard.

Precondition: Admin adds new places information in the system. The admin replies to the users questions.

Postcondition: Admin views the messages from users on the admin dashboard. Admin can add new places .

Comment/reviewed SSD

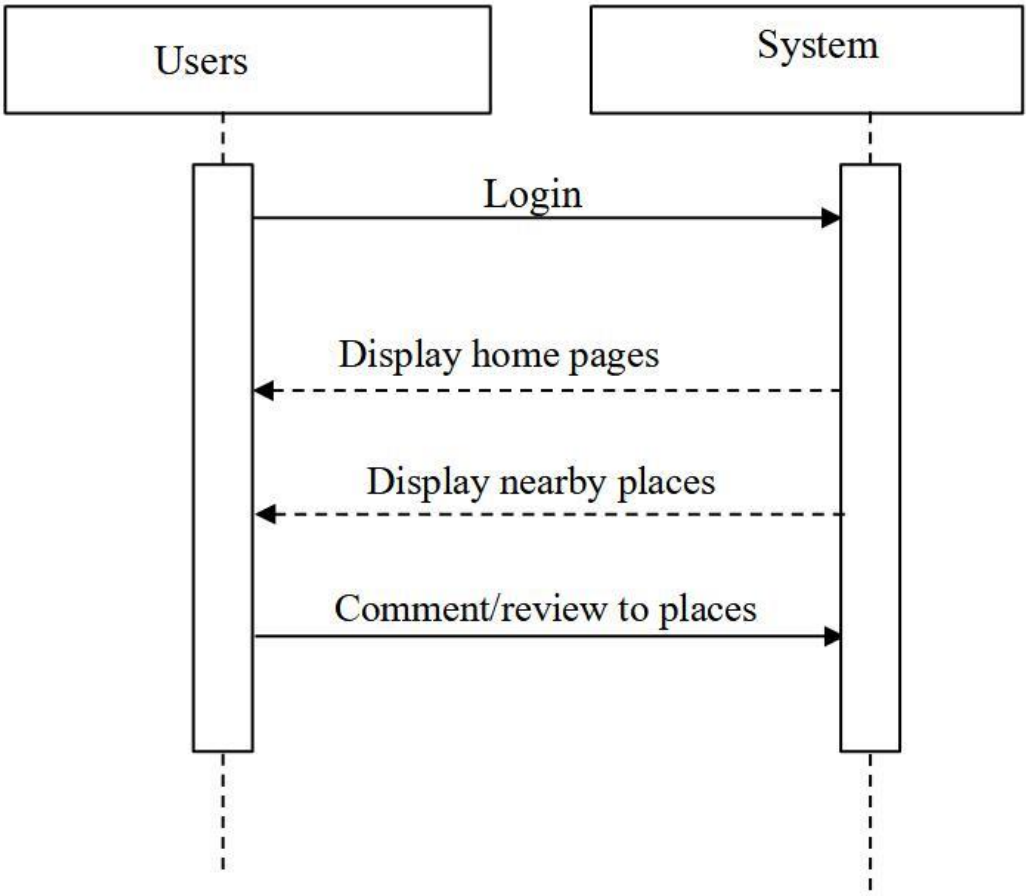


Fig: Comment/review

Name: Comment

Actor: Users

Description: The User can comment and write reviews on the visited places.

Precondition: The Users get the comment options on every place.

Postcondition: The users can write comments and also reply to the specific comments.

Sequence Diagram, ER Diagram and Activity Diagram:

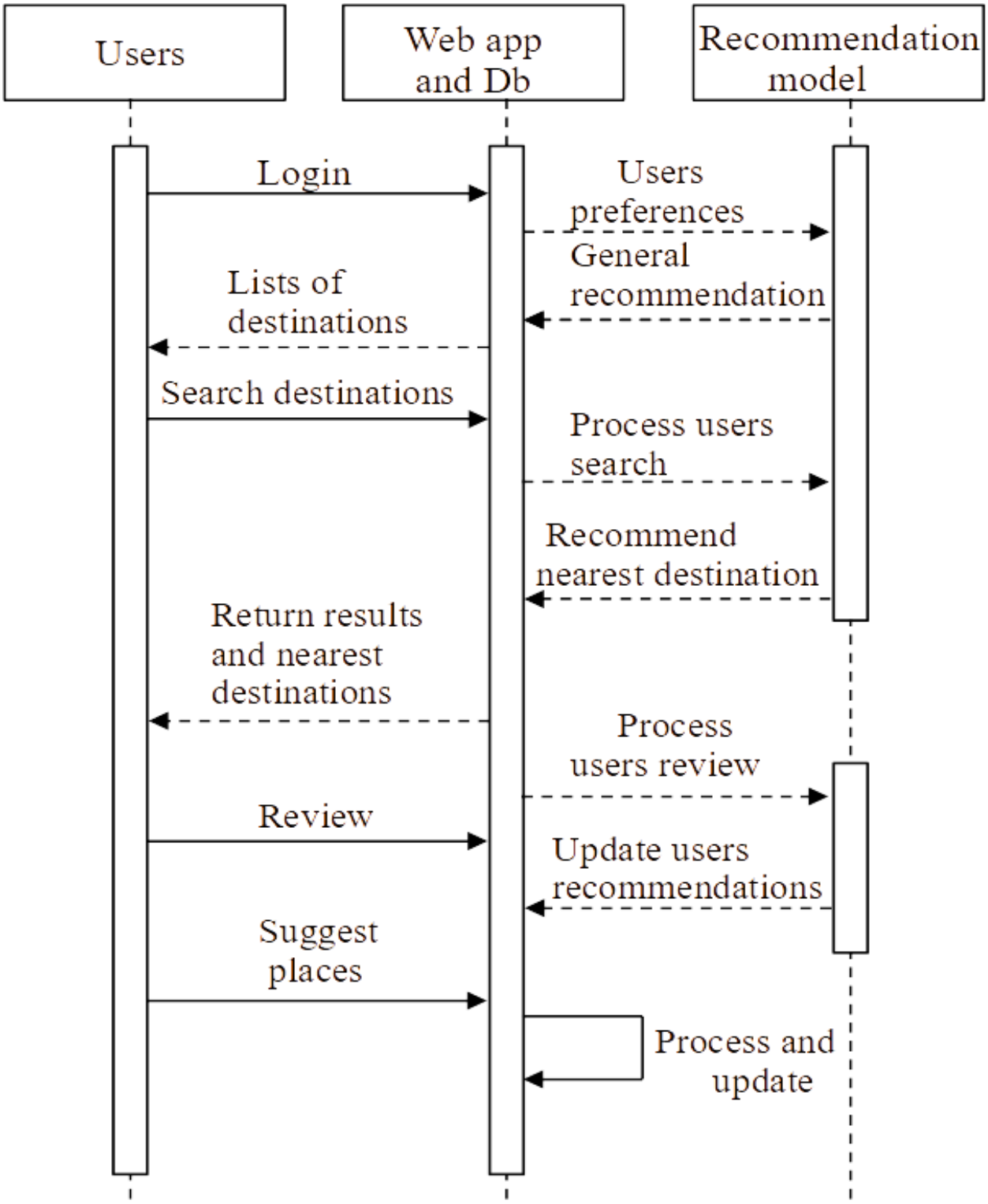


Fig: Sequence Diagram

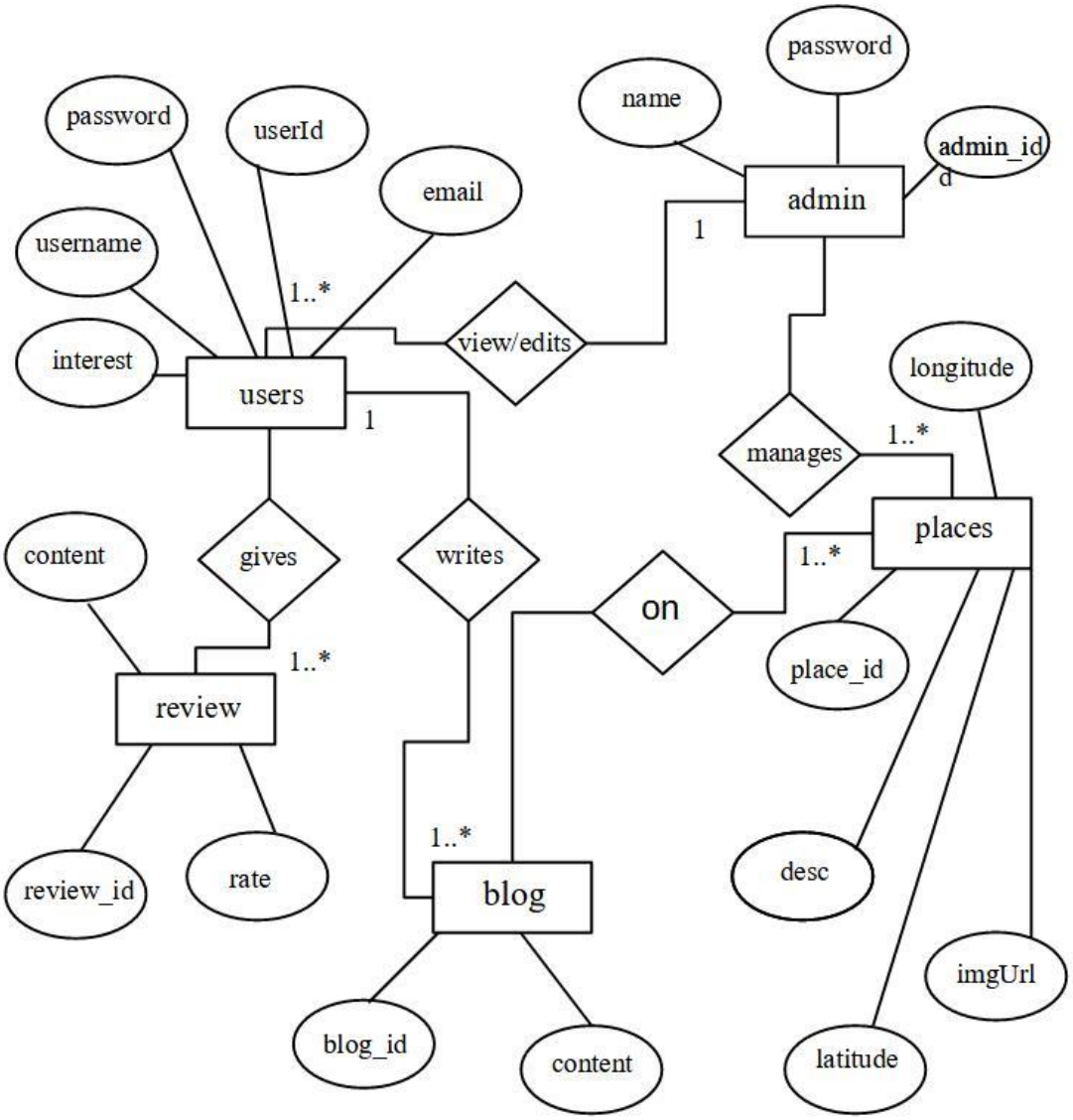


Fig: ER Diagram OF Discovering Nepal

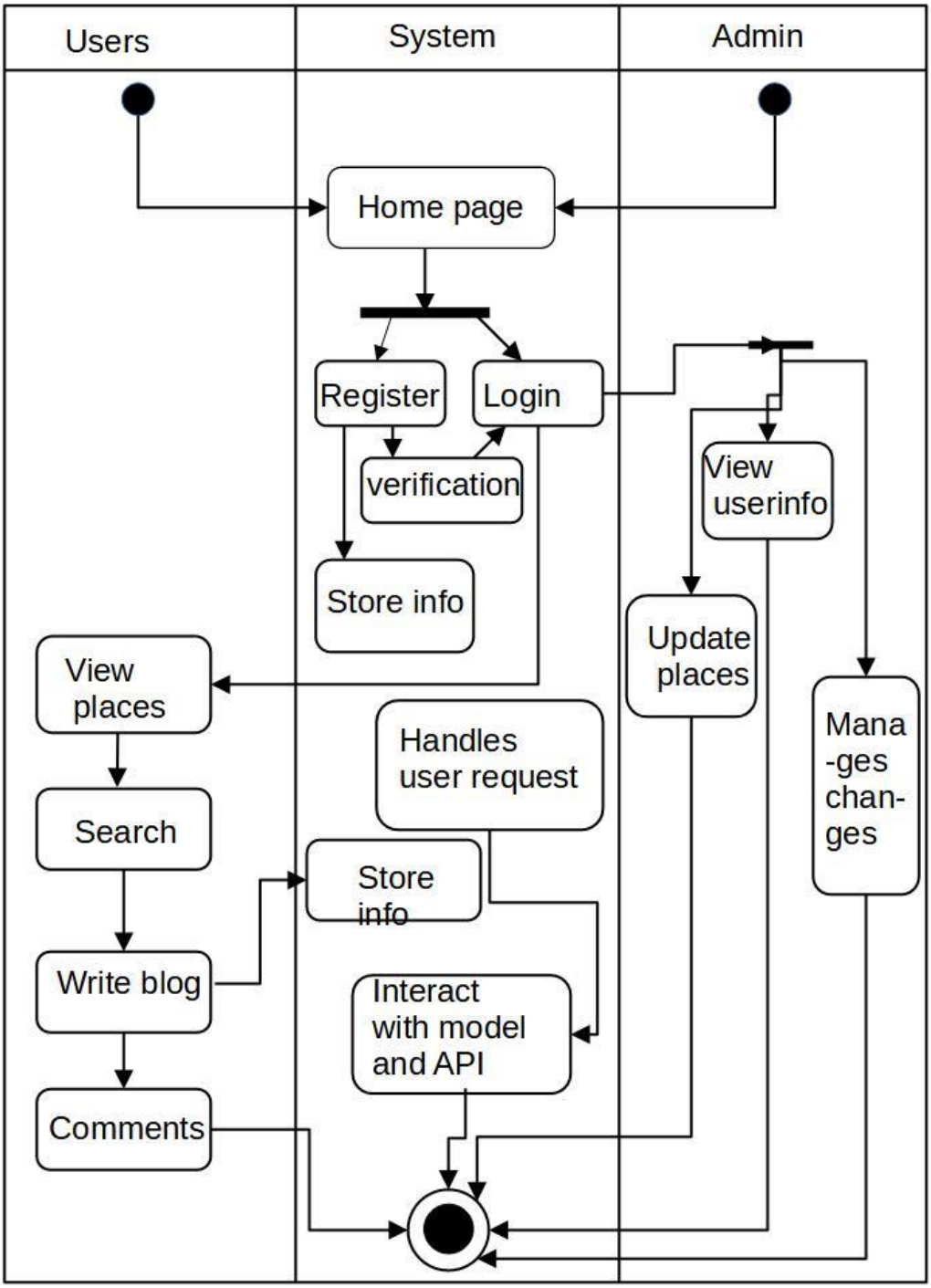


fig: Activity diagram

PROPOSED DELIVERABLES

Deliverables is concerned with the detailed features to be included in the software that solve the problems mentioned in the description of the problem area section.

The "core" features are as follows:

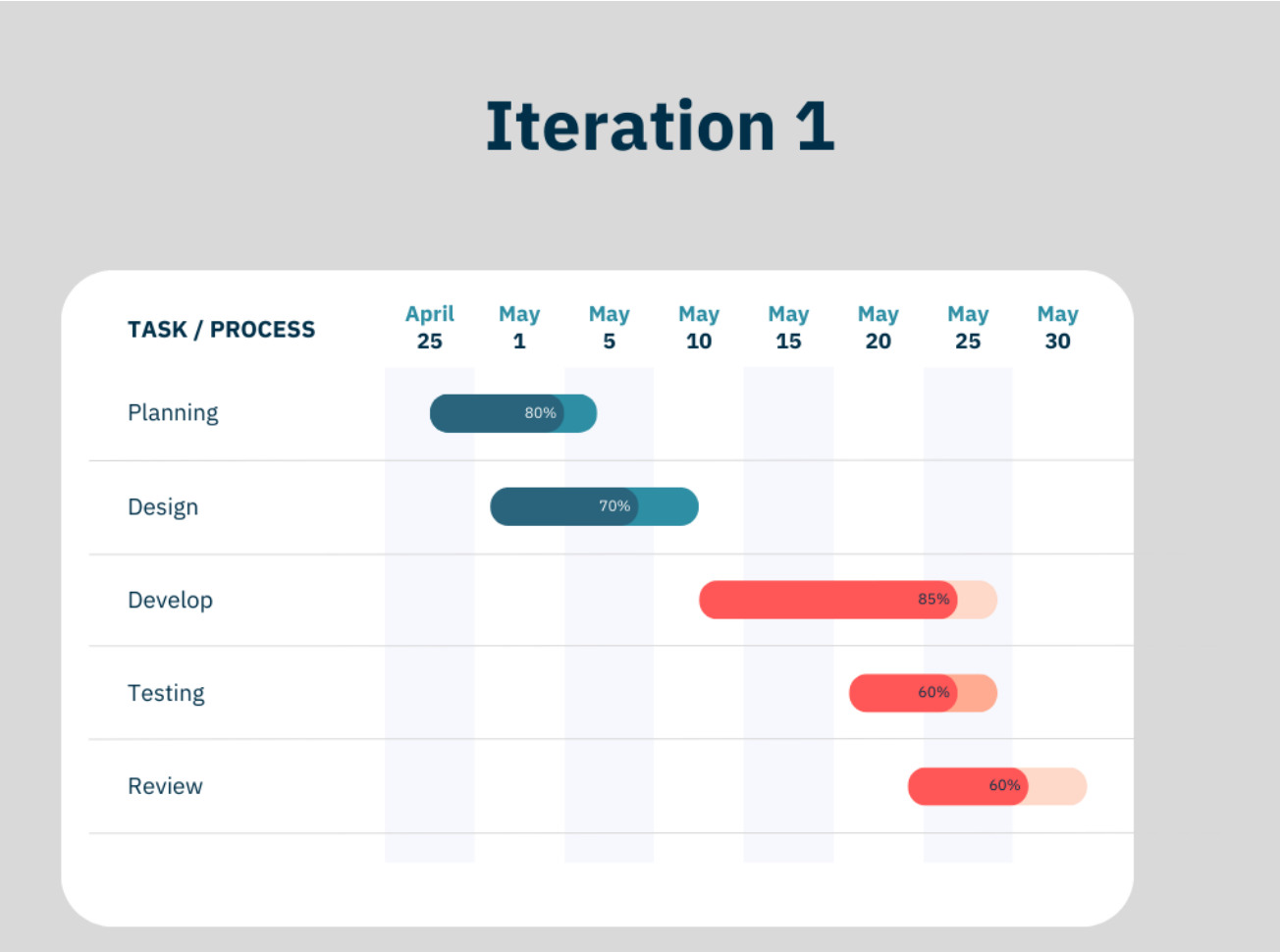
- Provide recommendations on the basis of user preference.
- Suggests nearby not so popular places while the user views any places.
- Estimate extra expenses while visiting nearby places from actual recommended places.
- Ask users to suggest new places to enlist in the system.
- Display the route to the destination from the user's current location.
- Provide ratings and reviews as feedback to improve the system.

PROJECT TASK AND TIME SCHEDULE

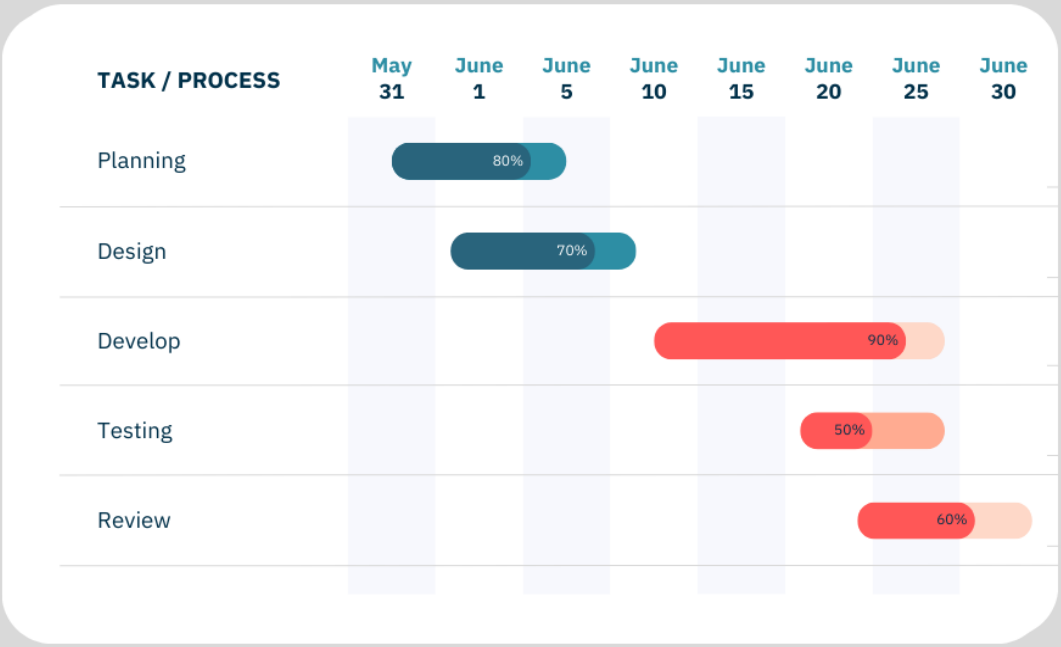
Work Assignment:

Team-members name	Planning	Design	Front-end	Backend	Testing
Aayush	✓	✓		✓	✓
Ashotoush	✓	✓	✓		✓
Rabin	✓	✓		✓	✓

Time schedule:



Iteration 2



Iteration 3

Bibliography:

[1] " KNN Algorithm" retrieved from https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm [Accessed May 02 2023]

[2] "Brief on Recommender system" retrieved from https://towardsdatascience.com/brief-on-recommender-systems-b86a1068a4dd [Accessed May 07 2023]

[3] "How to Build a Recommendation System in Python" retrieved from https://365datascience.com/tutorials/how-to-build-recommendation-system-in-python [Accessed May 07 2023]

[4] "Building a Recommender System from Scratch" retrieved from https://www.jillcates.com/pydata-workshop/html/tutorial.html [Accessed May 09 2023]

[5] "What is Agile Methodology?" retrieved from What Is Agile Methodology? (A Beginner's Guide) [2023] • Asana [Accessed May 10 2023]