VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

CSE3026– NoSQL Databases

J Component- Project Report

Review I

<u>Project Title</u>: <The Future of Travel Recommendation with MongoDB and Aviation API>

By

20MIA1088	***Aditya Kumar Singh***
20MIA1164	***Sakshi Chauhan***

M.Tech CSE with Specialization in Business Analytics

Submitted to

Dr.A.Bhuvaneswari,
Assistant Professor Senior,
SCOPE, VIT, Chennai

School of Computer Science and Engineering

September 2023

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

School of Computing Science and Engineering

VIT Chennai

Vandalur - Kelambakkam Road, Chennai - 600 127

Fall Sem 22-23

Worklet details

Programme	*M.Tech with Specialization in Business Analytics*	
Course Name / Code	*NoSQL Databases/ CSE3086*	
Slot	*C2 slot*	
Faculty Name	*Dr.A.Bhuvaneswari*	
Digital Assignment	*1*	
Team Members Name | Reg. No	*Aditya Kumar Singh*	*20MIA1088*
	Sakshi Chauhan	*20MIA1164*

Team Members(s) Contributions – Tentatively planned for implementation:

Worklet Tasks	*Contributor's Names*
Dataset Collection	*Aditya kumar Singh, Sakshi Chauhan*
Preprocessing	*Aditya kumar Singh*
Architecture/ Model/ Flow diagram	*Aditya kumar Singh, Sakshi Chauhan*
Model building (suitable algorithm)	*Sakshi, Aditya kumar Singh*
Results – Tables, Graphs	*Aditya kumar Singh*
Technical Report writing	*Aditya kumar Singh, Sakshi Chauhan*
Presentation preparation	*Aditya kumar Singh, Sakshi Chauhan*

ABSTRACT

The role of technology in enhancing the consumer experience has grown more significant in the constantly evolving world of travel and tourism. In order to reinvent the future of travel recommendations, this project explores the combination of NoSQL databases, particularly MongoDB, with Aviation APIs. Industry leaders in online travel booking platforms Expedia and Kayak are excellent examples of how ongoing innovation in this area is changing the way individuals plan and book their trips. The significance of MongoDB and Aviation APIs in transforming travel recommendation systems is explored in this abstract, highlighting how these technologies support the seamless travel experiences provided by industry leaders Expedia and Kayak. The Aviation API provides access to flight data from major airlines. This data can be used to power travel recommendations, such as finding the cheapest flights to a particular destination or finding flights that fit a traveller's specific needs. In this project, we'll use MongoDB to store and query flight data using the Aviation API to get real-time flight data. The system allows users to specify their own preferences for recommended trip destinations, and it is also highly flexible. Expedia and Kayak are pioneers in integrating MongoDB, a popular NoSQL database, into the development of dependable and scalable travel recommendation systems. MongoDB is a great option for storing an enormous amount of travel-related data because of its adaptability and capacity for handling unstructured data, including hotel listings, flight information, and customer reviews. Expedia and Kayak can effectively handle and retrieve various data types by utilizing MongoDB's document-oriented database format, ensuring that customers receive relevant and customized recommendations based on their preferences. In order to provide travellers with accurate, complete, and real-time information, MongoDB and

Aviation APIs are crucial. These APIs give users access to a wide range of aviation information, such as flight times, costs, and availability, as well as real-time flight monitoring. Travellers are able to make educated decisions when booking, rescheduling, or making last-minute adjustments to their travel plans thanks to Expedia's and Kayak's use of Aviation APIs to keep them up to date on the latest developments about their flights. Travelers are empowered by this incorporation of real-time data, which gives them the resources they need to easily negotiate the challenges of air travel. In order to provide highly personalized travel recommendations, online travel booking platforms analyze customer preferences and behaviours using machine learning and artificial intelligence algorithms. These platforms can make recommendations for trips, lodging, and flights based on user data mining, making the decision-making process easier and increasing consumer happiness. A working prototype of a recommendation system for travel that makes use of MongoDB and the Aviation API is what is anticipated to come out of this project. Performance and scalability will be assessed for the prototype. The prototype will be improved using the evaluation's findings, and suggestions for additional study in this area will also be made.

INTRODUCTION

The travel and tourism sector has undergone a significant transition recently, with technology increasingly influencing how travellers plan and book their trips. The fusion of NoSQL databases, particularly MongoDB, and Aviation APIs offers the potential to completely transform the way that travel suggestions are made in this age of constant innovation. In order to improve the travel booking experience for customers, this project intends to investigate the intersection of these technologies and their applications.

PROBLEM STATEMENT

Nowadays travellers have an overwhelming variety of options to choose from while arranging travel plans, including options for flights and lodging. It is difficult to give customers personalized recommendations that are appropriate for their needs and situations while also taking into account their preferences, spending limits, and changing conditions in real-time. Traditional relational databases find it challenging to manage the complexity and size of travel data, which is distinguished by its variety and unstructured nature. Another difficult task in this industry where punctuality is essential is to keep passengers informed about the most recent flight information, costs, and availability.

OBJECTIVES

The main goal of this project is to create a cutting-edge travel recommendation system that makes use of Aviation APIs and the flexible NoSQL database MongoDB. With personalized, real-time, and data-driven recommendations for flights, hotels, and locations, this technology will give travellers more power. Particularly, the ***objectives are as follows****:*

*1.****MongoDB integration:*** *Using the document-oriented database model of MongoDB to effectively store and manage a vast collection of travel-related data, such as airline information, hotel listings, and travel preferences.*

2. ***Using aviation APIs*** *to obtain real-time flight information, timetables, rates, and availability allows travellers to make well-informed decisions and changes to their travel arrangements.*

3. ***Offering highly personalised travel recommendations*** *to speed up the decision-making process, machine learning and artificial intelligence algorithms are being developed to analyze user behaviour and interests.*

4. ***Performance and Scalability****: The travel recommendation system concept is being tested for performance and scalability to make sure it can manage the needs of a large number of users.*

CHALLENGES

Travel data includes details like hotel descriptions, customer ratings, airline schedules, and more and is by nature diverse and unstructured. Although MongoDB's document-oriented database format is ideally suited for managing unstructured data, it might be difficult to create a database schema that is effective and can handle this variety. The difficulty is in arranging the data while preserving MongoDB's flexibility in a way that makes easy retrieval and analysis possible. It takes a lot of work to create recommendation systems that offer practical and relevant travel suggestions. It requires knowledge of data science, machine learning, and user behaviour. The difficulty comes in developing algorithms that consider a wide range of factors, including user preferences, previous booking history, financial limits, and real-time flight data. As they gain knowledge from user interactions, these algorithms must constantly change and advance to remain accurate and useful for travellers. Important security and privacy issues arise while handling user data in a travel recommendation system. When making reservations, travellers reveal sensitive information, such as personal information and payment data. It is crucial to ensure the security of sensitive data against any breaches.

LITERATURE SURVEY

S.no.	***Title***	***Author/ University/Journal***	***Technique***	***Result***
1.	***Recommending Customized Trips Based on the Combination of Travel Regions***	***Wolfgang Wörndl, Michael Ludwig, and Daniel Herzog*** ***Technische Universität München, Germany***	*The authors of the paper applied the* ***knapsack problem technique*** *to recommend travel regions for independent travellers. The knapsack problem is an optimization problem in which we are given a set of items, each with a weight and a value, and a knapsack with a limited capacity. The goal is to find the subset of items that has the maximum total value and that fits into the knapsack.*	*The best results were given by the authors' algorithm, which used the* ***knapsack problem technique****. The algorithm was able to generate recommendations that were more relevant and feasible for the users than the baseline algorithm. The algorithm was also robust to changes in the user's budget and the popularity of the travel regions.*
2.	***A Survey of Travel Recommender System***	***Roopesh L R, Tulasi. B*** ***Christ University***	*The technique used in this paper is* ***Hybrid Filtering****, which combines multiple recommendation algorithms such as* ***Collaborative Filtering, Demographic Filtering, Content-Based Filtering, and K-NN algorithms*** *This approach is used to overcome the cold-start issues and*	***Hybrid recommender systems have outperformed content-based and collaborative filtering systems*** *and have been given the best results. This is because hybrid systems can combine the strengths of*

			providing personalized recommendations in the travel and tourism domain	*both content-based and collaborative filtering systems.*
3.	***A Review on Travel Recommendation Techniques***	***P Sushmita Singh*** ***International Journal of Scientific & Engineering Research***	***<u>There are several travel recommendation techniques used in this paper</u>****:* *(i).* ***<u>Based on Cloud Computing and HADOOP</u>****: An architecture called TouchMap is proposed for travel recommendation. This architecture utilizes the HADOOP framework and MapReduce technique to process a large amount of point of interest (POI) data related to different attractions via parallel computing. The architecture consists of two major functional components: the POI search component and the trip planning component.* *(ii).* ***<u>Based on Hybridization Filtering</u>****: A Package Attraction-based Trip Recommender (PATR) framework was developed. This framework combines tour packages and*	***Collaborative filtering has been shown to be effective in travel recommendation systems. This is because collaborative filtering systems can use the ratings of other users to recommend items to the current user.***

			user-preferred points of interest (POIs) to accurately recommend a trip. It utilizes a Score Inference Model to obtain scores for attractions and tour packages based on user choices, and a Hybrid Trip Mine algorithm to generate the best possible trip by efficiently considering POIs and tour packages.	
4.	***A Thematic Travel Recommendation System Using an Augmented Big Data Analytical Model***	***Suriya Priya R. Asaithambi, Ramanathan Venkatraman and Sitalakshmi Venkatraman*** ***National University of Singapore***	*The technique used in this paper is a combination of* ***big data processing, machine learning, sentiment mining, time series forecasting, and artificial neural networks*** *to build an effective recommendation system for tourism activities.*	***Image-based prediction models performed well and achieved a high 97% accuracy based on the validation datasets.***
5.	***Research on Personalized Recommender System for Tourism Information Service***	***Huang Yu, Yao Dan, Luo Jing, Zhang Mu***	***The techniques used in this paper*** *include* ***content-based retrieval, the Apriori algorithm, statistical technology based on Bayesian classification, and the Weka system for data analysis.*** *These techniques are used for personalized recommendations in the travel recommender system.*	***The Apriori algorithm*** *gives the best results for personalized recommendations in tourism information services.*

6.	***Design and Implementation of a Personalized Tourism Recommendation System Based on the Data Mining and Collaborative Filtering Algorithm***	***Xiang Nan, Kayo kanato and Xiaolan Wang*** ***Kyoto University of Information Technology, Japan***	***<u>The techniques used in this paper include</u>*** *collaborative filtering, data mining, machine learning, support vector regression (SVR), and multilayer perceptron (MLP) models.*	*Support vector regression (SVR) and multilayer perceptron (MLP) models give the best results in this paper.*
7.	***Intelligent travel recommendation system by mining attributes from community-contributed photos***	***Subramaniyaswamy V, Vijayakumar V, Logesh R and Indragandhi V***	***<u>The techniques used in this paper include</u>*** *1. Mining GPS trajectory data to extract interesting locations and travel sequences,* *2. Analyzing travelogues (blogs) to obtain trip-related knowledge* *3. Systematically adopting large-scale photo databases to discover important landmarks.* *4. Extracting people attributes such as gender, race, age, and travel season from community-contributed photos.* *5. Determining the correlation between people attributes and likely travel locations using mutual information and* *6. Building a travel recommendation*	***The results show that there is a significant correlation between people attributes and travel location predictions.***

			model assisted by a Probabilistic *Bayesian model.*	
8.	***Personalized Travel Recommendation Systems: A Study of Machine Learning Approaches in Tourism***	***Mohamed BADOUCH, Mehdi BOUTAOUNTE*** ***Ibn Zohr University, Agadir, Morocco***	***The techniques used in this paper include*** *machine learning algorithms, collaborative filtering, content-based filtering, hybrid recommender systems, Markov models, and neural networks.*	***Hybrid recommender systems*** *have been* ***outperformed*** *and given the* ***best results*** *to* ***content-based and collaborative filtering systems****.*
9.	***A Collaborative Location-Based Travel Recommendation System through Enhanced Rating Prediction for the Group of Users***	***Logesh Ravi and Subramaniyaswamy Vairavasundaram*** ***SASTRA University, Thanjavur***	***The research paper discusses*** *various techniques used in recommendation for tourism like* ***Enhanced Rating Prediction Algorithm****. Some of the techniques mentioned in the paper include* ***collaborative filtering, content-based filtering, knowledge-based filtering, and social filtering****. Additionally, heuristic-based approaches, multi-agent systems, and AI techniques are also utilized in travel*	***The Enhanced Rating Prediction (ERP) algorithm gives the best results****. The ERP algorithm was able to improve the accuracy of recommendations by up to 10% compared to other rating prediction algorithms.*

			recommendation systems.	
10.	***Smart Tourism Recommendation Model: A Systematic Literature Review***	***Agung Trisetyarso, Edi Abdurachman and Yaya Heryadi*** ***Bina Nusantara University, Jakarta, Indonesia***	***The paper mentions several methods used in the research like Smart Tourism, Digital Tourism, Travel Recommendation Systems etc.***	***Collaborative filtering gives accuracy in the range of 0.76 - 0.91,*** ***Content-based filtering gives accuracy in the range of 0.66 - 0.86,*** ***Hybrid recommender system gives accuracy in the range 0.77 - 0.92.***

DATASET AND DATABASE SPECIFIC TOOL TO BE USED

For our project, we will need a dataset of flight information. We will use the Aviation ***API from the AviationStack Website****. An aviation API is a software intermediary that allows two applications to talk to each other. In the context of our project, the aviation API will allow us to access flight data from various sources.*

The aviation API from aviationstack.com provides a variety of flight data, including*:*

- *Flight status: The current status of a flight, such as "on time" or "delayed".*
- *Flight schedules: The scheduled departure and arrival times for a flight.*
- *Airport information: The location, runways, and terminals of an airport.*
- *Aircraft information: The type, capacity, and range of an aircraft.*

To collect data using the aviation API from aviationstack.com, we need to perform these steps*:*

- *Signing up for a free API key.*

- *Using the API key to authenticate our requests.*
- *Sending a request to the API endpoint for the data that we want to collect.*

***The API will return the requested data in JSON format**. We can store this data in a NoSQL database, such as MongoDB. The aviation API can play a valuable role in our project by providing us with access to real-time flight data. This data can be used to build travel recommendations that are personalized to the specific needs of our users.*

Advantages of using Aviation API to collect real-time flight data

1. We can use the API to track the status of flights that our users are interested in. This can help us to provide them with up-to-date information about their flights and to notify them of any delays or cancellations.

2. We can use the API to recommend flights to our users based on their travel preferences. For example, we can recommend flights that are on time, have a good safety record, and are operated by a particular airline.

3. We can use the API to generate reports on flight data. These reports could be used to track trends in flight delays, cancellations, or fuel consumption.

The specific database tools that we will need to use for collecting and the analysis that we want to perform:

***1. MongoDB Compass**: This is a graphical user interface for MongoDB that can be used to manage databases, collections, and documents.*

*2. **MongoShell in Terminal**.*

DATASET DESCRIPTION

The Aviation API will collect a comprehensive collection of travel-related information. This can encompass various data sources, such as flight details, airline information, user preferences, travel reviews, and geographical data. Key attributes in the dataset include flight departure and arrival times, airline routes, aircraft types, user ratings and reviews, historical weather data for travel

destinations, user profiles with travel preferences like budget, destination type (beach, city, adventure, etc.), and travel history.

DATABASE: MongoDB

MongoDB is an excellent choice for a NoSQL database for our travel recommendation project because of its flexibility and scalability. We can structure our data in a JSON-like format, making it easy to accommodate various data types and complex relationships. MongoDB's document-oriented architecture is particularly suitable for handling unstructured and semi-structured data commonly found in travel datasets. We can create collections for airlines, flights, user profiles, reviews, and destinations. MongoDB's support for geospatial queries will also be beneficial for location-based recommendations.

AVIATION API

To enhance our dataset with real-time aviation information, we can integrate an aviation API. Some popular aviation APIs provide data on flight schedules, live flight tracking, airport details, and airline information. These APIs can be used to update our database with the latest flight data, including delays, cancellations, and gate changes. This real-time data can greatly improve the accuracy of our travel recommendations, allowing users to make informed decisions based on current flight statuses and availability.

ALGORITHMS/TECHNIQUE DESCRIPTION

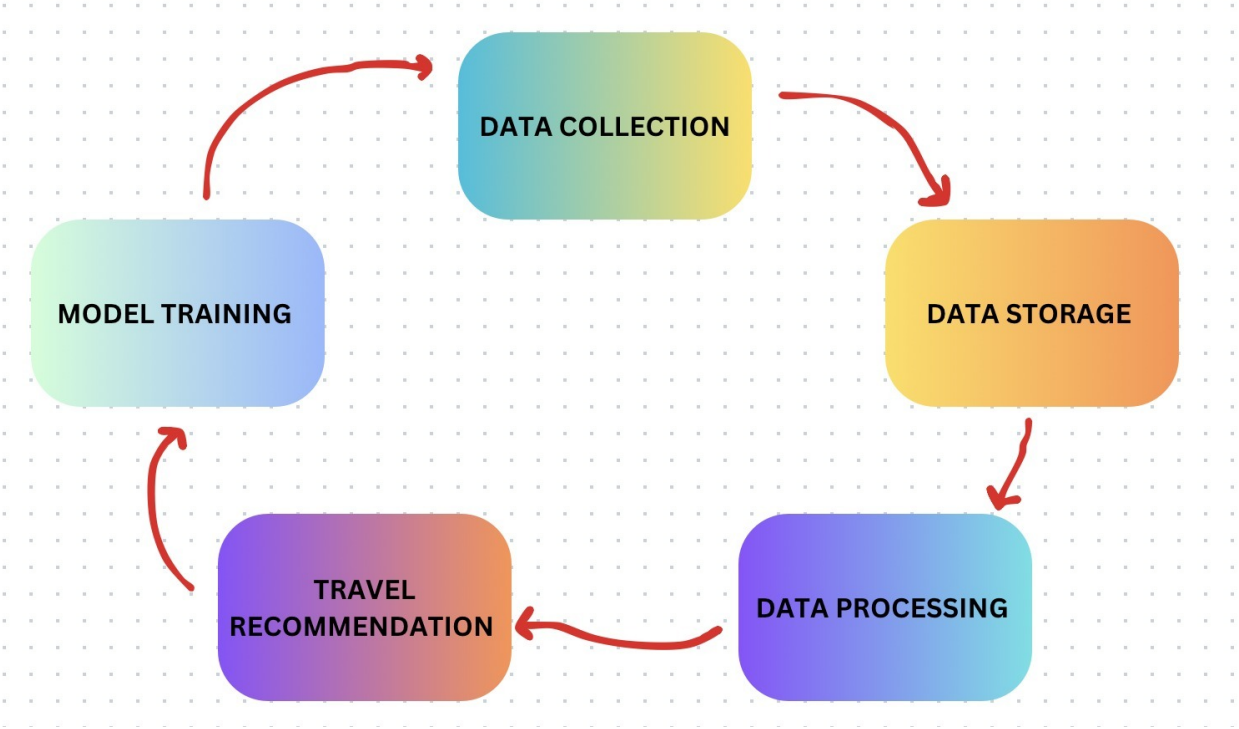
In this project we will be applying ***Hybrid filtering recommendation techniques****. Hybrid recommender systems combine two or more recommendation methods to gain better performance. Most commonly, we will use* ***collaborative filtering which is combined with a content-based filtering technique****. One common way to combine* ***collaborative filtering*** *and* ***content-based filtering*** *is to use a* ***weighted approach****. In this approach, the recommendations from the collaborative filtering algorithm are weighted by the results of the content-based filtering algorithm. This ensures that the recommendations are not only based on the user's past behaviour but also on the content of the items.*

*Another common way to combine **collaborative filtering and content-based filtering** is to use a **sequential approach**. In this approach, the collaborative filtering algorithm is used to generate the first round of recommendations. The content-based filtering algorithm is then used to refine these recommendations. This ensures that the recommendations are both relevant to the user's interests and based on the user's past behaviour.*

***Some advantages of using hybrid recommender systems**:*

1. *They can overcome the limitations of individual recommendation methods. For example, collaborative filtering can be biased towards popular items, while content-based filtering can be biased towards items that are similar to items that the user has already rated.*
2. *They can be more accurate than individual recommendation methods. This is because they take into account both the user's past behaviour and the content of the items.*
3. *They can be more personalized than individual recommendation methods. This is because they can take into account the user's individual preferences.*

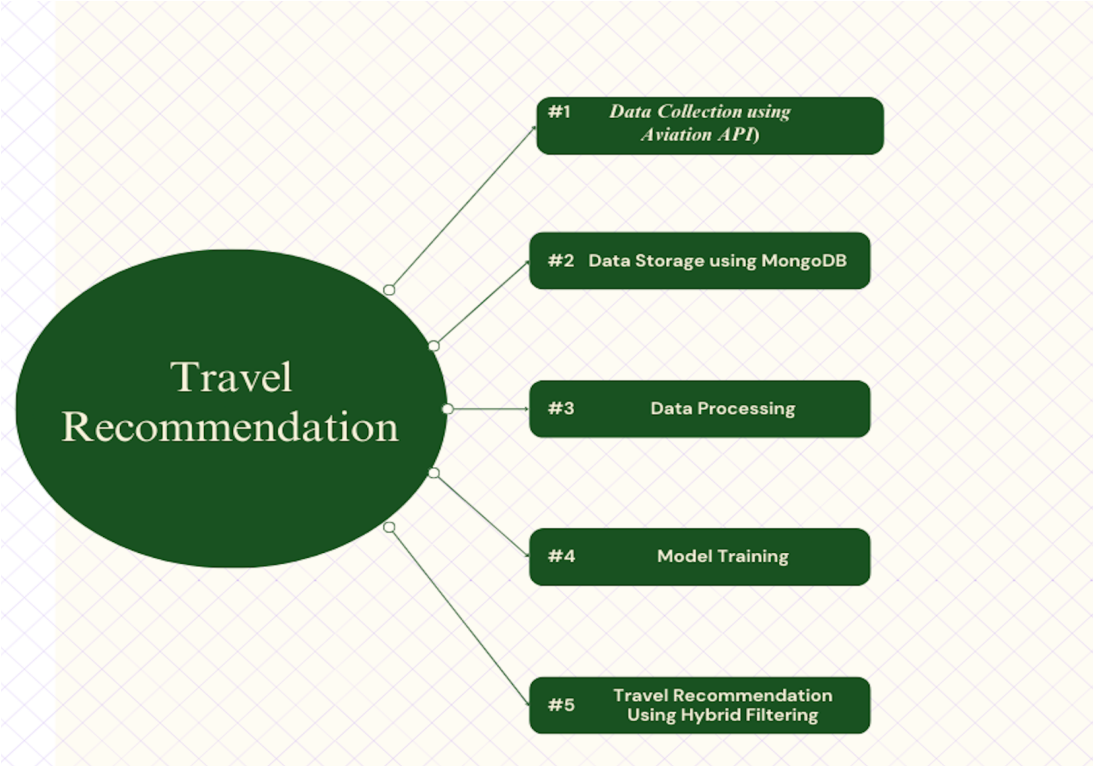
BLOCK DIAGRAM OF THE PROCESS FLOW OF THE PROJECT



Data Collection (Aviation API): *This component is responsible for gathering data from the Aviation API. It retrieves real-time flight information, airport data, airline details, and other aviation-related data.*

Data Storage (MongoDB): *MongoDB is used to store the collected data. It includes user profiles, preferences, historical travel data, and aviation data obtained from the API.*

Data Processing: *This component processes and analyzes data from MongoDB and the Aviation API. It involves data cleansing, transformation, and aggregation to prepare the data for model training and recommendation.*



Model Training: *In this step, machine learning models are trained using the processed data. These models can be recommendation algorithms, collaborative filtering models, or any other relevant machine learning models that learn from historical travel data and user preferences.*

Travel Recommendation: *Once the models are trained, this component generates personalized travel recommendations for users based on their preferences and historical data. It provides recommendations such as flight options, destinations, accommodations, and travel itineraries.*

GitHub Repository link

https://github.com/iamadi1709/The-Future-of-Travel-Recommendation-with-MongoDB-and-Aviation-API/blob/main/README.md

REFERENCES

[1] *Recommending Customized Trips Based on the Combination of Travel Regions Wolfgang Wörndl, Michael Ludwig, and Daniel Herzog Department of Computer Science Technische Universität München, Germany*

[2] *A Survey of Travel Recommender SystemRoopesh L R1*, Tulasi.B11 Dept. Computer Science, CHRIST Deemed to be University, Bengaluru, India*

[3] *International Journal of Scientific & Engineering Research, Volume 9, Issue 10, October-2018 ISSN 2229-5518 A Review on Travel Recommendation Techniques P Sushmita Singh*

[4] *A Thematic Travel Recommendation System Using an Augmented Big Data Analytical Model*

[5] *Research on Personalized Recommender System for Tourism Information Service Huang Yu Yao Dan Luo Jing Zhang Mu * Shenzhen Tourism College of Jinan University 6 QiaoCheng East Road, Overseas Chinese Town,Shenzhen 518053, Guangdong, China*

[6] *Design and Implementation of a Personalized Tourism Recommendation System Based on the Data Mining and Collaborative Filtering Algorithm*

[7] *Intelligent Travel Recommendation System by Mining Attributes from Community Contributed Photos December 2015Procedia Computer Science*

50:447-455 DOI:10.1016/j.procs.2015.04.014 LicenseCC BY-NC-ND 4.0 Authors: V. Subramaniyaswamy Vijayakumar Varadarajan Logesh R. V. Indragandhi VIT University

[8] Personalized Travel Recommendation Systems: A Study of Machine Learning Approaches in Tourism April 2023Shu Ju Cai Ji Yu Chu Li/Journal of Data Acquisition and Processing 38(02):867-877 Authors: Mohamed Badouch University Ibn Zohr - Agadir

[9] A Collaborative Location Based Travel Recommendation System through Enhanced Rating Prediction for the Group of Users Logesh Ravi 1 , Subramaniyaswamy Vairavasundaram 1 Affiliations expand PMID: 27069468 PMCID: PMC4812910 DOI: 10.1155/2016/1291358

[10] Smart Tourism Recommendation Model: A Systematic Literature Review Choirul Huda, Arief Ramadhan*, Agung Trisetyarso, Edi Abdurachman, Yaya Heryadi Computer Science Department, BINUS Graduate Program – Doctor of Computer Science Bina Nusantara University, Jakarta, Indonesia