A Smart Trip Advising System Based On Opinion Mining With Emoji Recognition

by

Mohammed Ziad Hassan
17101306
Mohammad Shahmidul Islam
17101096
Ahasanul Haque Abir
17101138
Anwara Iftesum Tanha
17101525
Akifa Jenia
17101513

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
January 2021

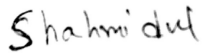
Declaration

It is hereby declared that

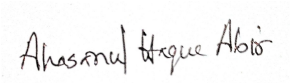
1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Mohammed Ziad Hassan
17101306



Mohammad Shahmidul Islam
17101096



Ahasanul Haque Abir
17101138

Anwara Iftesum Tanha
17101525

Akifa Jenia
17101513

Approval

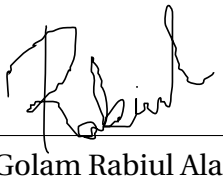
The thesis titled “A Smart Trip System Based On Traffic Prediction” submitted by

1. Mohammed Ziad Hassan (17101306)
2. Mohammad Shahmidul Islam (17101096)
3. Ahasanul Haque Abir (17101138)
4. Anwara Iftesum Tanha (17101525)
5. Akifa Jenia (17101513)

Of Fall, 2020 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on January 8, 2021.

Examining Committee:

Supervisor:
(Member)

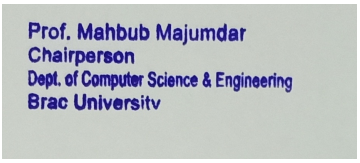


Md. Golam Rabiul Alam,PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Member)

Md. Golam Rabiul Alam
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)



Mahbubul Alam Majumdar, PhD
Professor and Dean
School of Data and Science
Department of Computer Science and Engineering
Brac University

Abstract

Opinion mining is an interesting research field that is used for analyzing the opinion of people about any product, place, and so on. Because based on people's opinion other people can make decisions about what they should do and what not; as well as other organizations can improve themselves. In our project, we will suggest a smart trip advising system where people can find out their desired places according to their budget and preferable season. We will collect reviews and opinions on which we will do opinion mining. Then we will also do an emoji recognition process to recognize the emoji which also contains people's opinions. After that, we will do sentiment analysis to analyze the data whether the opinions are positive or negative. Depending on those analyses we will rank the places and suggest to our users which place is suitable for them.

Keywords: Opinion mining, Emoji recognition, Sentiment analysis, Smart trip advising system.

Dedication (Optional)

Our research is dedicated to those people who want to have a refreshing and exciting trip. Here we are trying to provide something relevant to trip suggestion. Based on this we are trying to work with Bangladeshis tourism first. We want to make the process of finding places for trips simpler as well as getting feedback based on the opinion of users. We hope our research will make people happier and help them to have a wonderful trip.

Acknowledgement

First of all, we would like to thank Almighty Allah for helping us and showing us the right path as well. Then we would like to express our earnest and cordial gratitude towards our respected thesis supervisor Md. Golam Rabiul Alam, Associate Professor, Department of Computer Science and Engineering, BRAC University who helps us to improvise our ideas, gives guidelines and helpful directions to implement our thoughts. To all those who have supported us or at least wanted to support us, we take this sincere opportunity to express gratitude. We would also like to express our appreciation to our parents, family members, and friends for their kind encouragement and ongoing support for our work. Finally, we want to offer a feeling of thanks to anyone who supports us directly or indirectly and is interested in our work.

Table of Contents

List of Figures

List of Tables

List of Acronyms

The next list describes several symbols & abbreviation that will be later used within the body of the document

ML	Machine Learning
NLP	Natural Language Process
OP	Opinion Mining
POI	Point of Interest
POS	Parts of Speech

Chapter 1

Introduction

In order to earn a livelihood, to get a better life, we can see that people are always in a hurry, busy, and impatient. But these kinds of monotonous life can harm the mental peace and happiness of people. People cannot escape from their problems and responsibilities rather they have to face these things recklessly. Stepping away from the workplace and daily life is good for both mental and physical health. It emphasizes the strength and capacity to deal with future challenges. For this reason, people need some trip in order to make themselves happy, to get relief, to get new experience, new skills, and new knowledge to deal with the reality courageously. When a person decides to go for a trip he/she has to think about their budget or has to search for such kinds of places where he/she can go with his/her family or friends or by himself/herself as well. On account of this, he/she might have to inquiry on different online sites or different persons to get some review so that he/she can pre-plan his/her trip. Based on people's reviews through an online site or personally people decide which place is suitable according to their demands.

Human life is full of emotions and opinions. These emotions and opinions create a great impact on other people's decisions and activities. People's opinion helps a lot to make important decisions as well as to improvise a person or organization. But gathering opinion or review manually from other persons or online sites are really a very hectic process. That's why Opinion Mining(OM) becomes popular day by day. Opinion Mining helps to give an overall opinion by extracting necessary data from a huge number of data. After that, another most popular algorithm is Sentiment Analysis. Sentiment Analysis algorithm is used for getting the information about how positive or negative or neutral the opinions are. It helps to rank up several places or organizations by using people's opinions or reviews. Emoji recognition is another most interesting algorithm which identifies the meaning of emoji and uses it for getting some opinions because an emoji is not just an emoji it's a resource of opinion as well.

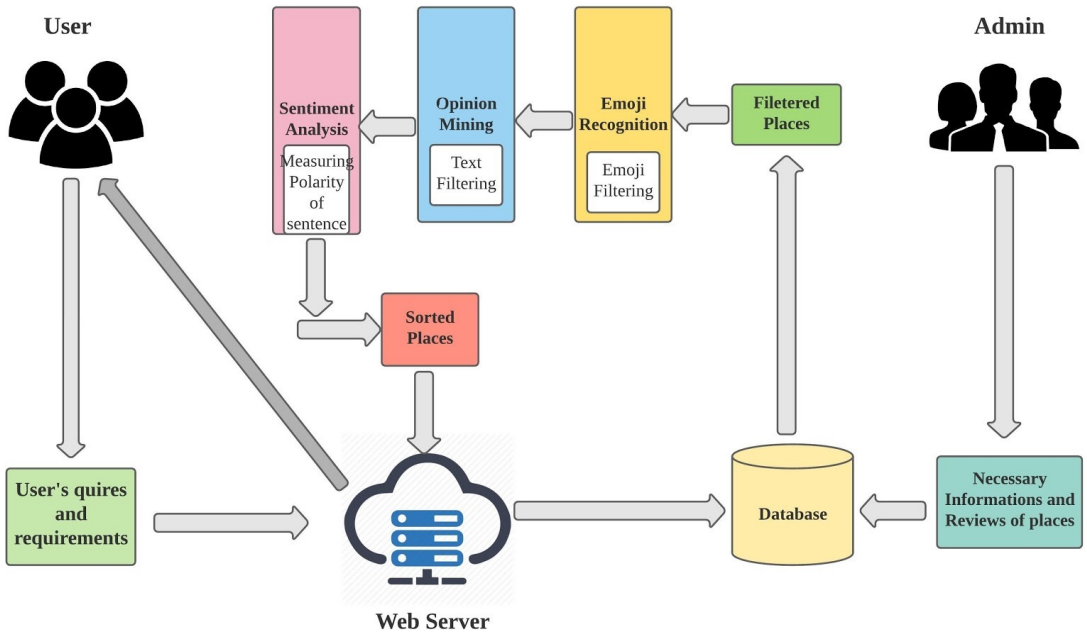


Figure 1.1: System Architecture of our project

By using all of this information, we are going to present a smart trip advising system where people can find out their desired place according to their demands for a trip; as well as get some sorts of places which are shown in the rank of people's opinion. We are trying to present a platform where users can search some places for a trip depending on their budget, season, and type of tour. Along with this, we are trying to show the result of search places on the rank of people's opinion on the same platform. We want to make people's life a little easier when they decide to get some refreshment and try to help them to make the best decision in order that at the end of the tour they will have full satisfaction with their decision.

1.1 Problem Statement

In past days travelers had to start planning for a long period. They had to search for all necessary information according to their preferences through the internet or people personally. Then gathering every possible review travelers used to make their trip plan according to other's reviews. That time travelers prepared a handy checklist for their desired tours. For a long period, they took the preparation for their upcoming tour which is a matter of a long time. But nowadays, people are getting busy from time to time in this modern period. So, in this thesis paper, we aim to create a single platform where our system can suggest the users according to their preferences. In one platform people can ask for a suggestion by providing their preferences such as budget, season, duration of the trip,trip members and so on. By filtering all the requirements our system will suggest the best option for the user whereas in past days people searched for information through various sources. That early day's process is time-consuming of course. So basically, our goal is to keep all of this information under one platform where our system will suggest the best option. This will be a great time saver and the

best solution for short time preparation takers. Thus, our system will help the travelers for their desired trip.

1.2 Research Objective

Our project intends to help travelers by suggesting to them the best place according to their preferences. Their preferences can be budget, tour duration, tour members, tour type, and so on. We are trying to build a system that can provide the best place for the user as well as provides some necessary information. Thus the user can get help from us and can plan accordingly. So our objective is to help users to find desirable places based on requirements, which may take a lot of time if he or she does it manually. This way our system not only saves users' time but also helps him or her to select the best place for a tour.
Nowadays people largely depend on others' opinions. People always think and want to know what others do and what others think about it. People want to do those things which have good responses from others. So we wanted to add others' opinions into our system. That's why we included opinion mining in our system. Our system not only will suggest places based on users requirements but also will consider other reviews to suggest. So our user will get both suggestions and others opinions to select his trip place.
Our system will work on some techniques. Firstly it will do opinion mining to analyze the texts. Then our system will do sentiment analysis to evaluate the collected reviews. Then our system will also recognize the emojis and it will be able to evaluate the emoji too. So the techniques we are using in our thesis paper -

- Opinion Mining
- Sentiment Analysis
- Emoji Recognition Evaluation

Our main objective is to save our user's valuable time. As in past days people had to search information through various sources and also from people personally. This process was time-consuming. So our system can help the travelers by proving all the information under one single platform. The user can now save their valuable time rather than wasting the tie by searching for information in various sources. Our research objective is to help travelers so that they can explore more places in their leisure time with less complexity.
As people are getting busy day by day so they have no time for taking a long time preparation and searching information in person or through the internet.
As people are getting busy day by day so they have no time for taking a long time preparation and searching information in person or through the internet.
In addition to that we will use contextual filtering in our trip system which would recommend the user location on a budget, and if it's a family or friends trip or not. The system will also give preference to the location that is suitable in the current season.

- Finds the best place for the travelers
- Provides approximate cost

- Provides approximate time duration
- Gathers all necessary information under one single platform
- Evaluates collected reviews
- Recognizes emojis and evaluates them too
- Suggest Place Based on Requirements and User Reviews

Chapter 2

Background Study

2.1 Background History

In our research we are considering the concept of machine learning, data science and NLP. The definitions and the short history of these concepts are given below:

An significant feature of current business and analysis is Machine Learning (ML). It is the analysis of computer algorithms that, over practice, improve automatically. Machine Learning is a branch of artificial intelligence that uses computational algorithms to learn from data and information autonomously. Computers in machine learning do not have to be directly programmed, but they can modify and upgrade their algorithms on their own. In order to make decisions without being directly trained to make those decisions, machine learning algorithms automatically construct a mathematical model using sample data, also known as 'training data. Machine learning algorithms currently allow machines to communicate with humans, drive cars autonomously, write and post sports match reports, and even find perpetrators of terrorism. [15], [27], [17]
The history of machine learning began with the first mathematical model of neural networks proposed by Walter Pitts and Warren McCulloch in 1943, in the scientific paper 'A logical calculus of the ideas immanent in nervous action.' [29] Then, the trend of evolving the idea of machine learning steadily began. The "Turing Test" was created by Alan Turing in 1950 to decide whether a machine had real intelligence. After that in 1952, Arthur Samuel wrote the first program for machine learning and the program was of the checkers game. Frank Rosenblatt created the first neural network for computers (the perceptron) in 1957, simulating the human brain's thought processes. The "nearest neighbor" algorithm was published in 1967, encouraging computers to start using very simple identification of patterns. The "Stanford Cart", which can handle obstacles on its own in a space, was invented by students at Stanford University in 1979. In 1981, the principle of Interpretation Based Learning (EBL) was proposed by Gerald Dejong, in which a program evaluates training data and generates a general law that can be implemented by discarding unimportant data. In 1985, Terry Sejnowski develops NetTalk, which learns the same way a baby does to say words. Study on machine learning in the 1990s moved from a knowledge-driven approach to a data-driven technique. In 1997, the world chess champion was defeated by IBM's Deep Blue. In 2006, the term "deep learning" was invented by Geoffrey Hinton to describe emerging algorithms that enable computers to "see" and discern objects

and text in pictures and videos. The Microsoft Kinect was able to track 20 human characteristics at a rate of 30 times per second in 2010, enabling users to communicate with the computer through motions and gestures. In 2011, IBM's Watson defeated Jeopardy's human rivals. Google Brain has been built along with this, and the deep neural network can learn to explore and categorize objects much as a cat does. In 2012, Google's X Lab developed a machine learning algorithm that can autonomously browse videos from YouTube to classify videos featuring cats. And in 2014, Facebook is creating DeepFace, a software algorithm that can identify or validate people at the same level as humans can on images. In 2015, Amazon will announce its own platform for machine learning. The Distributed Machine Learning Toolkit is also developed by Microsoft, which allows the successful dissemination of machine learning problems across many computers. Google's artificial intelligence algorithm defeated a professional player in the Chinese board game Go in 2016, which is considered the most challenging board game in the world and is sometimes more difficult than chess. [27]

Data science is an interdisciplinary method that derives expertise and perspectives from various structured and unstructured data using scientific techniques, procedures, algorithms and systems. The science of data is tied to data processing, artificial learning and big data.[18] In the world of Data Analysis, statistics and the use of mathematical models are profoundly rooted. Data Science began with statistics and has grown to incorporate concepts/practices such as Artificial Intelligence, Machine Learning, the Internet of Things, etc. There has been an influx of new knowledge, or Big Data, through the rise of the Internet, the Internet of Things, and the rapid growth of data volumes accessible to companies. The application of Big Data began to extend to other areas, such as health, architecture, and social sciences, until the doors were opened by corporations aiming to maximize profitability and drive smarter decision-making. [19]
William S. Cleveland is often attributed to the modern conception of data science as a separate discipline. [31] In a 2001 paper, he proposed an extension of statistics into technical fields beyond theory; since this would radically shift the discipline, a new name was warranted. [28] In the next two years, "Data science" became more widely used: in 2002, the Data Science Journal was introduced by the Committee on Data for Science and Technology. Columbia University introduced The Journal of Data Science in 2003. [28] The Section on Statistical Learning and Data Mining of the American Statistical Association changed its name in 2014 to the Section on Statistical Learning and Data Science, reflecting the increasing importance of data science. [18]
The data scientist, DJ Patil and Jeff Hammerbacher are both frequently credited with popularizing the term "data science," but they also represent the contemporary data scientist, that is, one who extends his or her data-savvy information to any field that needs it, including healthcare, e-commerce, social media, and journalism just to name a few. Patil, the chief data scientist at the Science and Technology Policy Office of the United States, has a lengthy resume that includes LinkedIn, Greylock Partners, Skype, PayPal, and eBay stints. [30]

We have vast volumes of knowledge on many facets of our lives, and, at the same time, a surplus of cheap computing resources. Shopping, chatting, reading news, listening to music, looking for evidence, sharing our views, as most people know, all this is being monitored online. What we don't know is that, representing the shift in online

data collection, the "datafication" of our offline actions has already begun. Bring the two together, and there is a great deal to discover about our actions and who we are as a species, by definition. However, it's not just Internet numbers, it's banking, pharmaceuticals, bioinformatics, social care, government, education, shopping, the medical industry, and the list goes on. In most sectors and most industries, there is a rising impact of information. In certain cases, the volume of data gathered may be adequate to be called "big".
We're seeing the beginning of a huge feedback loop, culturally saturated, where our behavior affects the product and our behavior changes the product. This is made possible by technology: large-scale computer processing infrastructure, expanded memory, and bandwidth, as well as societal adoption of technology in the fabric of our lives. A decade ago, this wasn't true. [20] There is no consensus on the concept of data science yet and others think it a buzzword.

A subfield of linguistics, computer science and artificial intelligence, Natural Language Processing (NLP) is concerned with the relationships between computers and human language, in particular with how computers are trained to handle and interpret vast volumes of natural language data. The effect is a device that can "understanding" the content of texts, including the qualitative complexities of the vocabulary inside them. The program will then reliably collect details and ideas from the documents and categorize and arrange the documents themselves. [21]

In the mid-1930s, the first patents for "translating machines" were filed. One idea by Georges Artsrouni was clearly to use paper tape as an automated bilingual dictionary. More specific was the other plan, by Peter Troyanskii, a Russian. This included both the bilingual dictionary and, based on Esperanto, a way of dealing with grammatical roles between languages. [22]
Production of natural language had its origins in the 1950s. Already in 1950, Alan Turing published an article entitled "Computing Machinery and Intelligence" which suggested what is now called the Turing test as an intelligence criterion, a challenge involving the automatic analysis and development of natural language, but not formulated at the time as a problem separate from artificial intelligence. [21] In the early 1970s, NLP (Neuro-linguistic Programming) was first introduced in the U.S. to research the reasoning and behavioral skills employed by highly productive and active individuals. Millions of people around the world are now using it globally in such varied areas as administration, distribution, communications, public relations, education, counseling, military and police, sport, and so on. [23]
SHRDLU, a natural language system running in small "blocks worlds" with limited vocabulary, were several notably popular NLP systems developed in the 1960s. [22] An intense and productive partnership took place between 1972 and 1974, with Grinder benefiting from Bandler's understanding of psychotherapy and Grinder's experience of linguistics. [24] Many NLP systems were based on complicated collections of hand-written rules up until the 1980s. However, beginning in the late 1980s, with the advent of machine learning algorithms for language analysis, there was a revolution in NLP. Many of the notable early accomplishments occurred in the field of machine translation, primarily due to work at IBM Science, where more complex mathematical models were successively developed. [22]

2.2 An Overview of Existing Trip Advising System

Here are some of works done by different peoples on different time and different place. Some of them are mentioned below:
ATIPS: Automatic Travel Itinerary Planning System for Domestic Areas: They design an automatic travel itinerary planning system for the domestic area (ATIPS), basically peoples get themselves busy for making their trip plan it wastes their valuable time. They used an algorithm to minimize this time so that they can enjoy their valuable time. Simply by entering the travel time, the departure point, and the destination location, the system can automatically generate a travel itinerary. According to the results of their experiments, 70% of users reacted positively and were satisfied with their result. This result filled the requirement perfectly which reduce their planning time. And 82% of users were satisfied with the automatic user preference learning mechanism of ATIPS. Their algorithm also provides a framework which basically substitute previous system and processed and offers a very good process for making a travel planning. [25]
A Collaborative Location Based Travel Recommendation System through Enhanced Rating Prediction for the Group of Users:
In this process, they worked on social network data based on different kinds of recommender systems by considering usage of various recommendation algorithms, functionalities of systems, different types of interfaces, filtering techniques, and artificial intelligence techniques. They have also proposed a location recommendation system based on social pertinent trust walker (SPTW) and compared the results with the existing baseline random walk models. Later, they have enhanced the SPTW model for group of user's recommendations. The results obtained from the experiments have been satisfied every requirement. Using this kind of algorithm and process they sort out social network responses and then they recommend users expected result from the system. [9]
User-adapted travel planning system for personalized schedule recommendation:
They worked on very important and specific problems. One of them is people simply recommend the popular tracks most rather than the suitable one. Suppose someone is going to visit one place. If he ask anyone or any apps they will show him the popular one first rather than user requirements. And they can't even provide the plan schedule. To tackle these problems, they develop a personalized travel planning system that simultaneously considers all categories of user requirements and provides users with a travel schedule planning service that approximates automation. It will be provided totally as like the user wants. A novel travel schedule planning algorithm is embedded to plan travel schedules based on users' need. Even user can replace any unsatisfied travel unit from that interface. They have the feedback mechanism which will helps new users and provides a better accuracy rate. They take an experiment and result was outstanding. The results showed that participants who used the system with schedule planning have statistical significant on user satisfaction and use intention. They also analyzed the validity of applying this algorithm .This one was very satisfying to users. [5]

2.3 An Overview of Opinion Mining and Sentiment Analysis

Opinion Mining and Sentiment Analysis has been one of the most used and known branches of computer research of natural language processing (NLP). Opinion mining or emotion analysis is a technique of text analysis that uses computer linguistics and processing of natural language to automatically classify and derive thoughts or viewpoints from text (positive, negative, neutral, etc.). It helps a person to get into the minds of his/her clients and figure out what they want and hate, and why, so that he/she can build goods and services that suit their needs. A person can conduct opinion mining automatically, on almost any type of unstructured text, with very little human feedback required, when you have the right resources. Basically, opinion mining or sentiment analysis deals with the retrieval of evidence from the opinions of individuals. For a particular subject, it can gather results from the feelings or opinions of people. [25]
Although we know that both opinion mining and sentiment analysis are the same method, but between them there is a slide distinction. While sentiment analysis, a precursor to the field of opinion mining, explores how individuals feel about a given subject (whether positive or negative), opinion mining goes further to consider the drivers of discussion behind the sentiment. [19] For real-time data, sentiment analysis may process thousands of websites, tweets, emails, or polls in just minutes. Or you can do opinion mining over time and see how the classification of emotion increases or falls. Models for opinion mining and emotion analysis will concentrate on opinion polarity (positive, negative, neutral), personal emotions (angry, joyful, depressed, etc.), and motives or interests. This is opinion mining aimed at identifying and removing individual feelings from the text (anger, sadness, frustration, pleasure, etc.). Lexicons, or word lists identified by the emotions they denote, are used by some emotion recognition tools. On the other hand, sophisticated machine learning algorithms allow text analysis programs to learn from sample text directly, so that they can grasp the complexities of human language, even to the point of irony and sarcasm being found. [25]

2.4 An Overview of Emoji Recognition

Emoji are ideograms which are used in electronic messaging and web sites. In different genres emoji are existed including facial expressions, typical things, locations and weather forms, and animals. They are almost like emoticons, yet emoji are photographs rather than typographical approximations; in the strict context, the word "emoji" applies to images that can be interpreted as encoded characters, but by definition, it is often extended to messaging stickers.[20]

Originating on Japanese cell phones in 1997, after being introduced to many mobile operating systems, emojis became widely common worldwide in the 2010s. In the West, they are now known to be a major part of mainstream culture. The Face with Tears of Joy emoji was named the Word of the Year by Oxford Dictionaries in 2015. [26]
A simple text-based variant of the already developed Unicode emoji script, the emoji

was predated by the emoticon and possibly took influence from pictograms. Several efforts to modify the generic emoticon to make it more suitable for use were made in Europe, Japan, and the United States in the 1990s. The emoji is based on the premise of forming pictures using text markers. The release of Unicode's latest emojis in 2009 saw the debut of some of the most notable emojis used today. There were several teething problems with the release of the new emojis, with input from many on the cultural gaps between various countries and even misuse. In order to satisfy the needs of various communities, the success of emoji has caused pressure from manufacturers and foreign markets to introduce extra designs to the Unicode standard. Approximately 250 Emoji is added to Unicode 7.0.[26] These emojis contains human's emotions and intentions. By understanding these emotions and intentions through machine learning process we can get the sentimental result and get an overview from them.

Chapter 3

Literature Review

We have read several research papers and out of that seven papers had a lot of similarities with our topic which are reviewed below.

In the first paper we reviewed "Probabilistic Tourist Trip-planning with time-dependent human and Environment Factors", this paper proposed a probabilistic tourist trip-planning method considering time-dependent satisfaction and constraints. This satisfaction and constraints are modeled as random variables in the hybrid temporal Bayesian network. By using the Bayesian network, they can infer the probability of satisfaction and constraints. The next constraint genetic algorithm is employed to find the best itinerary/way by maximizing the expected satisfaction under probabilistic constraints. When tourists visit a city or region, they usually face the problem of deciding a visiting order among points of interest (POIs). This trip-planning problem involves several constraints such as the visiting time required for each POIs and budget limits. Here are some limitations of this research that the arrival time at a POI (Point of interest) cannot be deterministic and existing methods do not consider dynamic conditions of humans and environments such as tourists' fatigue and weather conditions. But in our project, we are dealing with the thing that suggests places to our user according to people's queries and showing results by using opinion mining and sentiment analysis. We are not dealing with the probability of time-dependent satisfaction and constraint. But the thought of people's satisfaction and constraints is going to help us in our project.[7]
In the second paper "Adaptive Trip Recommendation System: Balancing Travelers Among Point of Interests with MapReduce", the system recommends various information to the user which is based on their preferences, needs, or constraints. Besides this, the system also considers its constraints. The system is divided into two main components. They are offline analysis and recommendation engines. An offline analysis is analyzing past data. From these past data, the system sorts all information based on a user's point of interest. Then the system recreates a new data set like when, where, how people are interested to visit. Thus, gathering all past information this system recommends a user when, where, how they can visit. The recommendation engine has two stages. They are Pareto set computation and Stochastic optimal search. They both work parallelly in this system. Firstly, a user gives a query to the system. Then the Pareto set processes the user's query. Pareto set matches and compares them with the point of interest. After getting the output from the Pareto set the stochastic optimal search algorithm gives an optimal solution by applying the perturbation algorithm.

Finally, this output is sent to the user as a suggestion or recommendation.[12]
In the third researched paper named “Personalized Trip Planning by Integrating Multimodal User-Generated Content”, they build a multimodal website which is a merged version of multiple websites. They used the information from multiple websites that they need for their research and merged the data and used it on their website or webpage named “Photo Odyssey”. According to them, a photographer uses a lot of time in searching for places, photos, and camera equipment. Their website will give or show everything in one place to save time for the photographer. They used UGD (User Generated Data) like Text, Image, Numeric Value. They used a String Matching Algorithm to find different data sources online and presented that data in an ontology. They used 500PX, Flickr for photos, LonelyPlanet, TripAdvisor to find places or cities travel information, Photography Ontology for information about camera equipment and Google Map API for location. By using that information, they can give Map location of the city, Travel information and camera information to the user on their website “Photo Odyssey”. They also talked about getting information like place name, time of photo taken, used camera and lens, ISO other photography related information from picture description which they want to use on the website. They have filter options and customize options that give users full control of what they want to see or want to do. They faced the problem of integrating different types of data which they solved by an Ontology-based approach and using Interoperable Semantic Web Standards. To find relevant data and rank it according to merits they used String Matching Method and Contextual Filtering. This system is intended for the photographer where a photographer can get all information in one system.[6]
In the paper “T-DesP: Destination Prediction Based on Big Trajectory Data”, we found that whenever a destination is picked or visited with specific attributes, the recommendations are based on the location visited previously. However, there may be much passing by location across the route of the destination, which might be suitable as per attributes but is never recommended. Therefore, a novel model is proposed, T-Desp, to address this problem. To execute this model a method named Mirror Absorbing Markov chain. This model works by the map-matched trajectories into two parts one for destination and passing by location. The missing values in the tensors are filled through a tensor decomposition approach. Secondly, using the node pairs in the current time slot probability of changing time and trajectory is learned from the probability of routes. After that, absorbing tensors can be constructed to store the absorbing probability between node-pair in a particular time slot. This probability is used for each node query, rank them according to personalization, and predict the most probable destination. However, the paper falls behind predicting any destination during any festivals, also there is no prediction traffic factor, and there is no personalization for people. The solution for the festival can be categorizing every festival and finding a pattern for it. For considering traffic, the pattern of the traffic should be in consideration during the node probability calculation.[8]

In the fifth paper, we reviewed "A Spatio-temporal loyalty-based model", they tried to find the actual number of loyal passengers. Then they evaluated the data of loyal passengers to get the actual value of passenger pressure. The daily fixed passenger numbers are Loyal Passenger. load (Loyal Passenger Detection) where they divide every station as a zone to get a map. Every passenger has a specific id and how frequently they are purchasing their tickets are explaining where that person is loyal

or not, by analyzing those data they got the actual numbers of loyal passengers. MFCP (Multiple features combined prediction) is, after gathering all the data of loyal passengers, they add other data of time, location, and weather and finally they calculated and showed their result of the prediction. For error checking, they find the values of RMSE, MAE, Normalized MAE. Then they compared this model with SVR, BPNN, and GBDT and got the best value from loyal customers. Prediction Accuracy results are, STloyal was 90.92% correct for weekdays and 81.64% correct for the weekend. Among the four models, this loyalty-based model performed as the best traffic prediction model for the subway. [14]

In the sixth research paper, we reviewed "Opinion Mining and Sentiment Analysis –An Assessment of Peoples' Belief: A Survey", this paper is concentrating on the fundamental definitions of Opinion Mining, the study of the linguistic tools needed for Opinion Mining, the few techniques for ML-based on their use and meaning for the analysis, the assessment of classifications of Sentiment Analysis and its numerous applications. This paper has surveyed and met up with the field of Sentiment Analysis and Opinion Mining. This paper has some sections which have contained different types of discussions on basic information about OM and sentiment analysis, important ML techniques, evaluating sentiment analysis's classifications, discussion on its wide range of use or applications, different types of NLP tools which are being used for sentiment analysis. It is really very challenging to do a task with sentiment analysis. In our project we are dealing with opinion mining as well as sentiment analysis also; so this paper helps us a lot to gather the necessary information for our project.[3]

In the seventh paper which is named "Mining opinion components from unstructured reviews: A review", this paper provides a comprehensive literature analysis about the computational methods, models, and algorithms for opinion mining components from unstructured feedback. The purpose of this paper is to provide the new knowledge of opinion mining to the researchers and students while they frame new theories and improve the process further. This paper is also organized in different sections which contain a general summary of mining topics, their implementations, and associated fields of opinion. Technical perspectives of opinion mining based on opinion components are also explained in a section, other sections provide summaries of opinions, offer an outline of problems and concerns as well. This analysis reveals that substantial focus has been paid to opinion mining in the last few years, focused on online content and the developments in Web 2.0 technologies. In order to analyze opinion representation, opinion mining models, opinion elements, and related issues, this analysis utilizes social networks and web blogs, the most popularly used sources for opinion recovery. There is a detailed discussion of a variety of computational models and linguistic characteristics relevant to opinion mining, component analysis, and opinion-target recognition.[4]

In the eighth paper which is named as "Opinion Mining and Sentiment Analysis"; this paper covers methods and methods that promise to enable opinion-oriented information-seeking structures directly. In contrast with those already present in more conventional fact-based analysis, this paper's attention is on approaches that aim to solve the new problems faced by sentimental technological applications. This paper provides material on the overview of the evaluative text and on wider privacy, exploitation and economic effect problems that are generated by the creation of opinion-oriented information-access services. A discussion of available tools, benchmark databases, and appraisal campaigns is also presented to encourage potential

work. This paper discussed about an additional audience for applications capable of measuring customer opinion automatically, as reflected in no small part in online venues, are businesses keen to learn how they interpret their goods and services.[1]

In the ninth paper which is named as "Facial Emoji Recognition", this paper provides information about that in the development of the human computer interface, facial expression detection and interpretation has gained great interest as it offers a normal and powerful way to interact between people. Facial emoji recognizer is an end-user app that recognizes the person's face in the video taken by the camera. On the frame, which varies with the shift in emotions, the smiley related to the voice of the person in the video is displayed. In human speech and relationships, facial expressions are important. They are also used as an important instrument in behavioral research and in medical fields. The goal of this paper was to create an intelligent facial expression classification method using the CNN algorithm. The Haar classifier is used to detect the face and the CNN algorithm is used to detect the expression and to send the emoticon appropriate to the expression as the output. A human emotion detector using emoticon using machine learning, python to predict people's emotions and represent them using emoticon, is proposed in this paper. Live videos taken from the webcam are the subject of this paper's method. [16]

In the tenth paper, which name is "Emotional Recognition Using Facial Expression by Emoji in Real Time", this paper focuses on the assistance of pictorial representations of facial expressions, it reflects on the role of emoji in promoting emotional identification. The pictorial portrayal of the facial expression of human beings is Emoji's. In the facilitation of individual subjective interactions, they are very helpful. This research work explores emotional recognition using real-time emoji facial expression. In addition, it further develops the criteria of facial expression assessment and real-time perception of facial emotion recognition. The created application contains six human expressions, including emotions that are neutral, anxiety, rage, happy, sad, and surprise. These phrases are the real expressions that human beings communicate. Because of their capacity to better communicate human emotions and the way they promote contact between individuals, the investigations of such speech are important.[10]

In the eleventh thesis paper which is named as, "Social Media Emoji Analysis, Correlations and Trust Modeling"; this paper discussed about the secret to social networking is how the mobile devices used by the majority of individuals are used to enable the fast sharing of media and news between people. Twitter is an ever-growing social media site where people share tweets, or small posts, to see and respond to all of their fans. While this study sample covers tweets over a period of 11 years, it is not an accurate measurement of how tweeting habits are influenced by incidents and different periods of the year. An overview of how the distribution of tweeted emojis can vary over various periods of time may provide a clearer explanation of users' state of mind during these periods of time. Using emoji analysis in tweets to spot sarcasm was quickly addressed but otherwise not commented upon. This may be done by analyzing the polarity of a tweet and matching it with the emoji's average polarity. In this article, though confidence modeling is discussed, an in-depth look at the development of feature vectors will offer a theoretically more detailed model.[13]

Chapter 4

Proposed System and Methodology

4.1 Proposed System

In our thesis paper, we have took two data sets and used all of them in our model. Firstly, we have collected the distance of each tourist attraction by placing each of the districts in the center. For example, if a user is going to select any place for a recommendation when he/she is in Dhaka. The distance given to the query will be based on that. Secondly, there is a data set where all the tourist locations are labeled with attributes like cost, type of trip like- friends or family, and the season, which is suitable to visit.
We have trained our model so that if any user gives input with their preferences then the system will give the best suitable place as an output.
The major attributes that we have used to build our system is the duration of the trip, approximate cost of the trip, type of trip, and season too. A description about the attributes of our system are given below:

- **Duration of the trip :** Our system has recommend a suitable place based on the travel time it requires to reach the destination otherwise, it will not be feasible for the user to plan for the trip. For example, if it's a 1-day trip it would be less than 120 hours of travel time. Similarly, for 2 days it would be less than 175 hours. However, if it's more than 3 days it would choose all possible values. For this reason, we have used this attribute in our system.

- **Type of trip:** A filter for choosing the type of the trip means the option for users to choose a place where he/she can go with his/her friends or family. This kind of requirement is also given in our system. This is necessary because some places are suitable for family trip. On the other side some are suitable for friendly trips. Some places are physically challenging where we cannot go with our family members but can go with our friends. For example, hiking in hilly terrain or exploring deep parts of the forest. This will not be suitable for children and women. So these types of places are not suitable for any family trip. That's why we have used this attribute.

- **Budget of the trip:** Everyone has a plan for the budget of any trip. It is quite natural that people who have a plan to complete the trip on a budget because of not having enough money. That's why we have used the attribute of budget in order

to make easier the process of finding out a place for a trip in a budget. Because some places are expensive to travel and some are comparatively cheaper. So our system will provide that facility to filter according to the user's amount of budget.

By using all of these information as an input in our system, our system will have done the contextual filtering process. After this filtering process, our system will get some places as an output. If we do not sort the places which are getting from the contextual filtering then our user will become confused because they would not understand that which place is most suitable and which is not so suitable according to their requirements. For this reason, we have used opinion mining and sentiment analysis process in order to rank the results of places based on people's opinion. We have collected people's reviews or opinion from different websites and apply the opinion mining and sentiment analysis algorithm for getting the polarity of those reviews. As we know that in the modern generation people love to use emoji in their statement and those emojis also carry some sorts of opinion; that's why we have also use the emoji recognition algorithm to analyze the emojis and getting a sentimental result from them. As opinion mining process cannot detect emojis so that we have used the emoji recognition process and detect its polarity. On the other hand, the opinion mining process have applying its algorithm on all the reviews of a place and the sentiment analysis process the polarity of the comments for a place. After that, both the polarity of emoji and sentences will be add with each other and give result of the overall polarity for a place. Thus, all the polarity of the filtering places will be occurred and after that based on those polarity a sorting algorithm will be applied on the filtered places. Then the sorted list of the filtered places will be shown in the display and a small description of the sorted places also shown in the display of a user.
Thus our system will help to save time as well as our system will serve a platform to the user where they can get places according to their requirements, able to know the most suitable place based on people's opinion as well as also able to know a small description of those places.

4.1.1 Constraints

Constraints play a vital role while designing and implementing a system. It shows how the system will behave or follow some fixed set of rules while giving an output. In our system, the season could be a factor while giving recommendations. Therefore, the tourist spots are most lucrative during the season so this comes first in the recommendation but if it's a budget constraint the system should show all locations on the possible budget.

4.2 Model Description

At first, the system will query the user about the user's requirements (duration, budget, types of tour, season). Then the system will take those inputs from the user and then send it to the server. The server will send it to the database where all sorts of data are stored. Applying algorithms all the information will be filtered according to the user's

requirements. After filtering, the database will give the output (selected tourist spot) to the server. Then the filtered places will go to the next procedure where those places will be sorted based on people's opinion. People's opinion will be separated into two parts these are - text and emoji; because the text identifier process cannot detect the emojis. So, the emoji recognizer will detect emoji and the opinion mining process will detect text as well as keep only necessary data which will be used for measuring polarity of sentence. After that, the sentiment analysis process will occur to measure the polarity of the sentence. Then based on those polarities filtered places will be sorted and sent to the server. After that, the server will send the sorted data to the user. Thus the whole procedure happened.

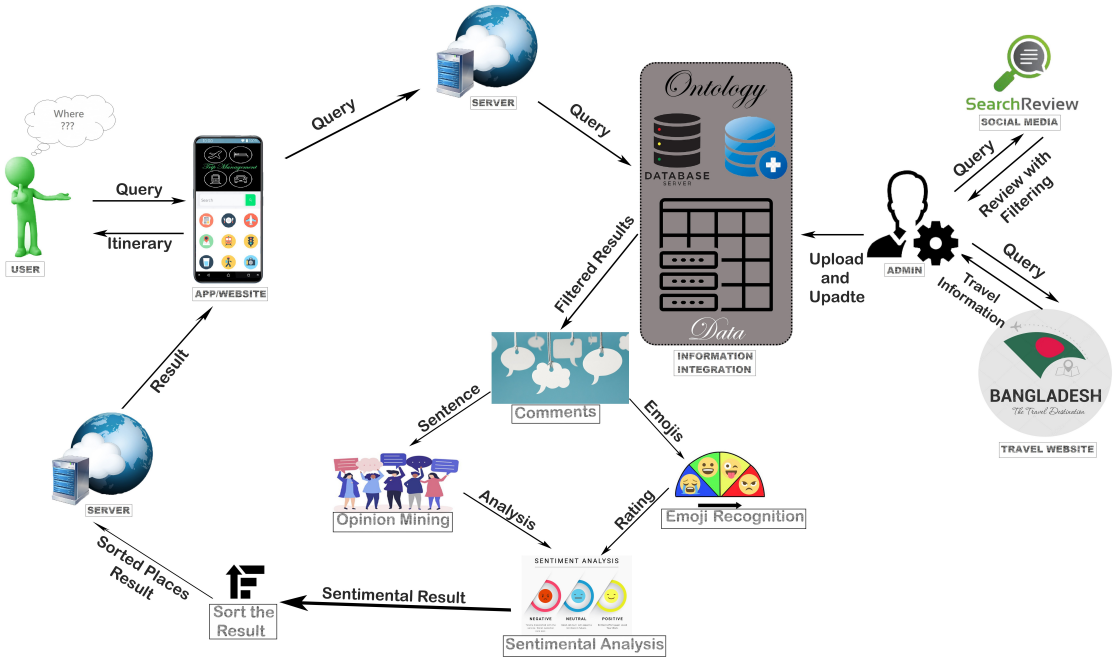


Figure 4.1: Model Diagram

On the other part, there is an admin who is controlling the database continuously when it is needed. The admin gathers information from social media, travel websites and so on. Then the admin filters all those information and reviews as well as keep only necessary information which is done manually. After that the admin updates and uploads the information in the database whenever it is needed.

4.3 Workflow Diagram

Here, Fig 4.2 shows a flowchart for the workflow of the system. The phases which are proposed in the system are followed the below mentioned steps:

- Firstly, the system will ask some questions from the user such as – duration of the trip, budget, season and so on.
- Our algorithm will apply on that information in order to filter the information from the database.

- Algorithm will be applied on the variable “Duration”. If the duration of the trip is longer/ the duration of the trip is shorter those requirements will go to the database and the system will go for the next variable.
- Next our algorithm will be applied on the variable “Budget”. If the budget is low/the budget is high those requirements will go to the database and the system will go for the next variable.

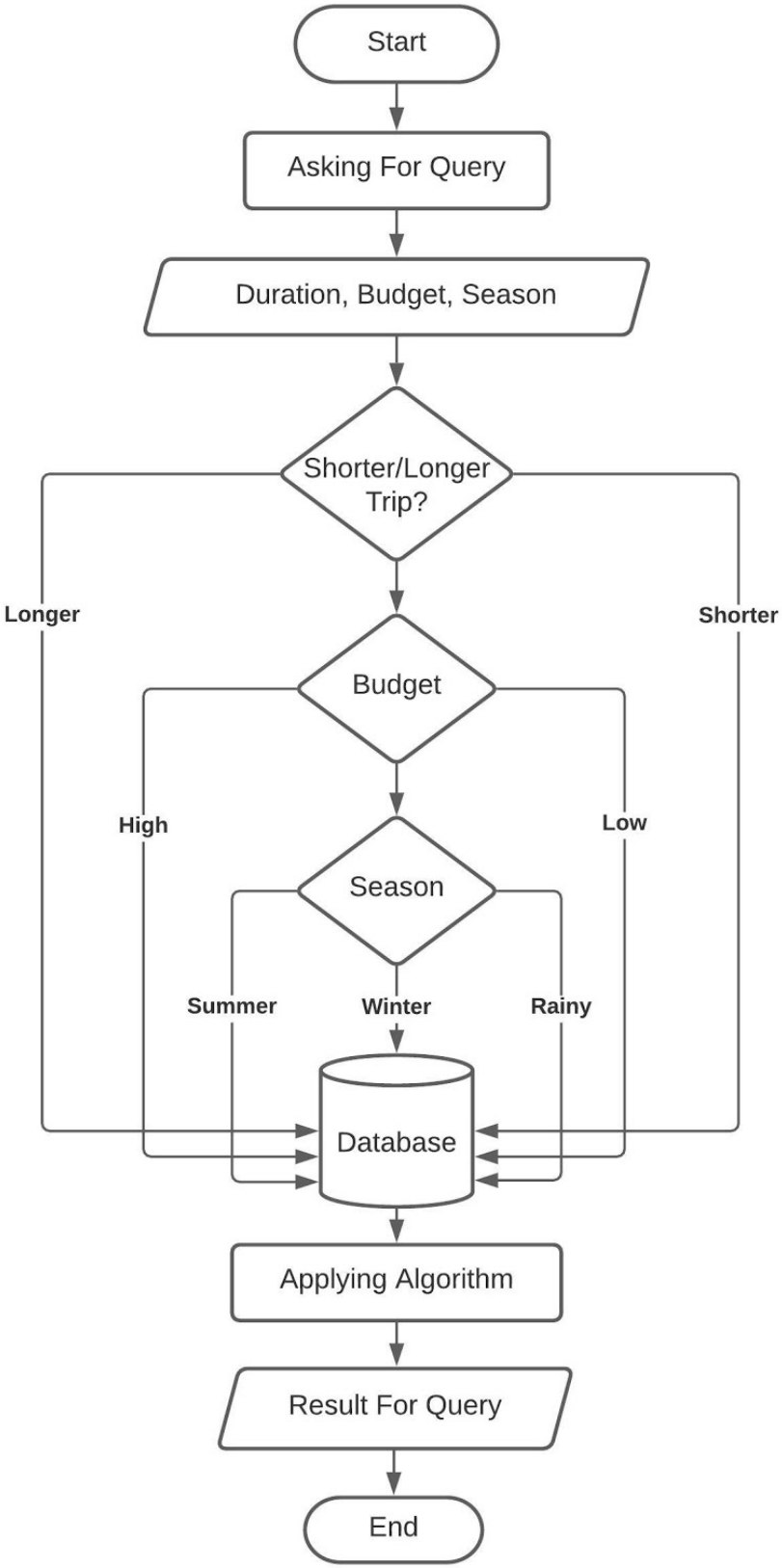


Figure 4.2: Flowchart of the workflow of User-Database System

- In the next variable “Season”, our algorithm will be applied on it to find out the desired season of the user. If a user chooses summer/winter/rainy season that

requirement will go to the database to find out the suitable place for the user.

- After filtering the information, this filtered information is being used in the database to extract data among lots of data.
- The extract data which is the suitable place for a trip will go to our ranking algorithm which is used for doing the sorting method by using opinion mining to understand which place is highly recommended.

After completing all these procedures, the suitable place along with its ranked up position will be shown to the user as an output.

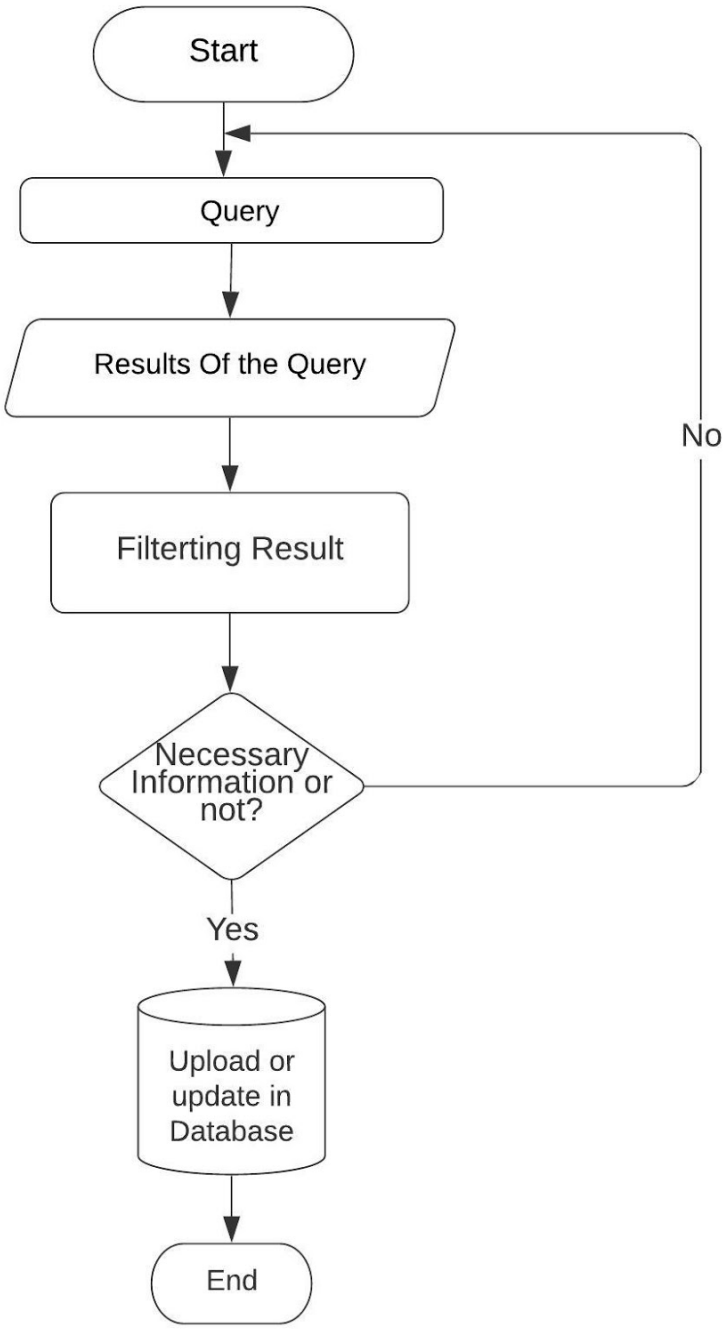


Figure 4.3: Flowchart of the workflow of Admin-Database System

- Admin will keep searching for his queries in different websites, social media, search reviews and so on.
- After searching, the admin will get some results of his queries.
- Then admin will filter that information manually and select which information are necessary and which are not. Along with this, the admin will put all the reviews in different folders to apply opinion mining, emoji recognition and sentiment analysis algorithms over those reviews.
- Then the necessary information will be uploaded and updated by the admin in the database and then he keeps searching until he gathers all his queries results.

After updating and uploading information in the database, the most important and interesting procedure will be applied to sort the places according to the concept of which place is most preferable for a user depending on people's review or opinion. Here in the Fig 4.3 shows the workflow of the raking system. The phases of the proposed system is given below:

- After getting the result of places which are getting through the filtering process by users information; the system will search reviews for each place.
- If the system will get several places as input for then the system will do the search review process until all the reviews of a certain place do not finish.
- Then the system will separate text and emoji because the opinion mining algorithm cannot recognize emojis.
- After separating text and emoji a condition will apply. If the separated part is emoji then the system will do an emoji recognition process and if the separated part is text then the system will do the opinion mining process.
- After that another condition will apply to analyze whether there is any review left or not. If that particular place has more reviews then the system will go through the search review process again and do the same procedure again. If that place does not have any review then the system will do the sentiment analysis process to get the polarity of those reviews.
- After doing the sentiment analysis process another condition will apply to find out that there is any place left or not. If the system has left any place then it will go to the system's input where the places are given as input. If the system does not have any more places then the system will sort the places by using the polarity of the reviews for each place which are getting through sentiment analysis. It describes which place is more preferable among those given places.
- Then finally the list of sorted places will be shown in the display.

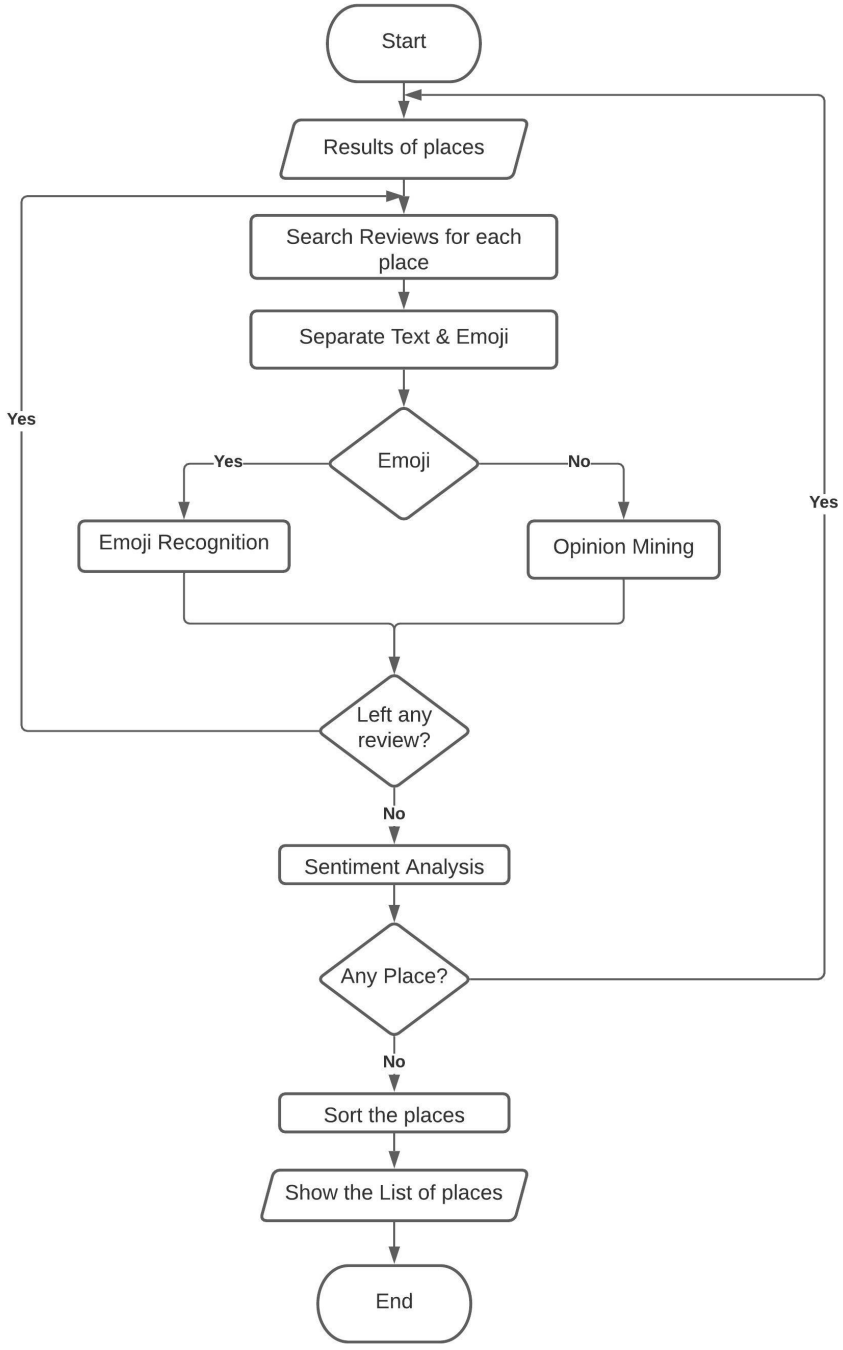


Figure 4.4: Flowchart of the workflow of Ranking System

Thus the whole process of our project "A Smart Trip Planning Based on Opinion Mining" has been completed. It will suggest places for trips based on users requirements and it will show those places which are ranking up based on people's review or opinion.

4.4 Decision Tree Based on Filtering Method

In our decision tree, the first decision is regarding the duration of the tour. Here after getting the input of duration it will search whether there is any tourist spot which can

be traveled within this duration or not. If there any tourist spot is found then it will go forward to another decision.

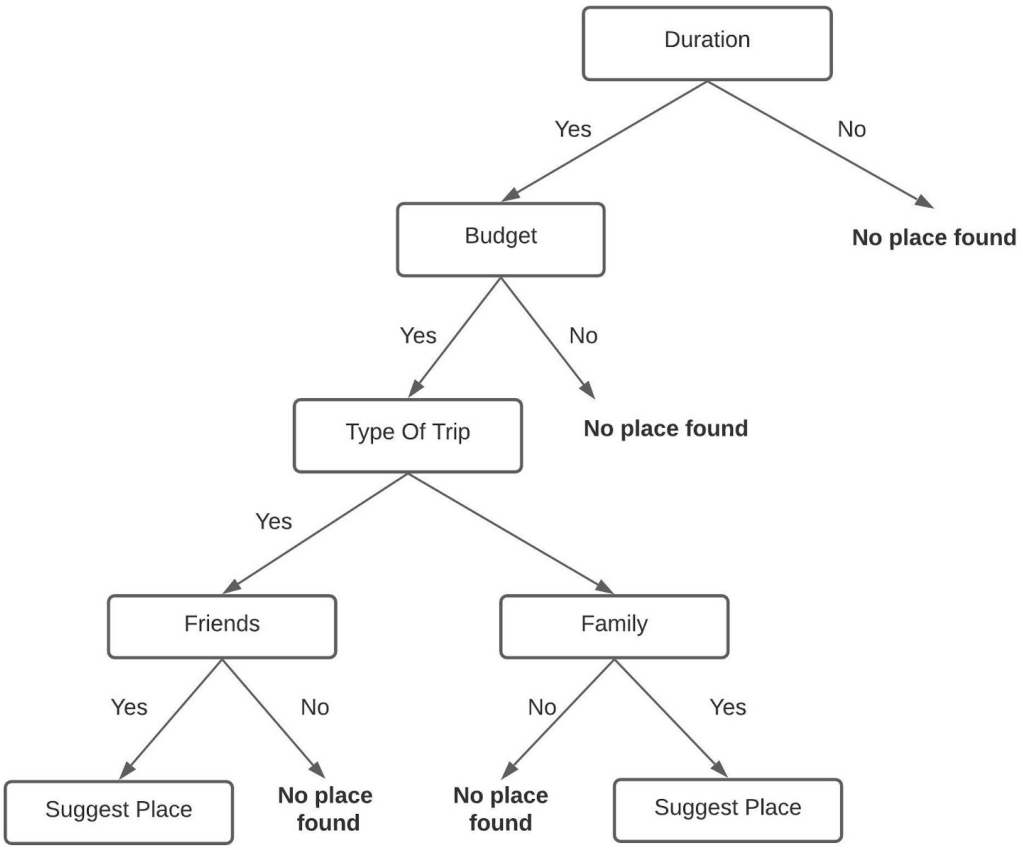


Figure 4.5: Decision Tree

On the other hand, if it does not find any tourist spot then it will show “No place found”. The second decision is regarding budget issues. In the same way, if any tourist spot is found which can be traveled within the given budget then it will move forward to another decision and if it does not find any tourist spot within the user’s budget then it will show “No place found”. The third and the last decision is regarding types of tour whether it is a family trip or friends’ trip. Then if any tourist spot is found by fulfilling all the requirements then it will suggest the tourist spot. On the other hand, if it does not find any tourist by fulfilling the requirements then it will show “No place found”. In this way, the whole decision tree will take all the decisions one by one.

Chapter 5

Data preliminary analysis and discussions on operations

5.1 Data Preliminary Analysis

Data collection is one of the most important phases for best outcome results. Moreover, it is a method of collecting all the qualitative and quantitative information on particular attributes for the purpose of evaluating the best result.

	Place	Dhaka	Bagerhat	Bandarbar	Barguna	Barisal	Bhola	Bogra	Brahmanb	Chandpur	Chittagong	Chuadanga	Comilla	Cox's Baza	Dinajpur	Faridpur	Feni	Gaibandha	Ga
2	Cox's Baza	391	512	121	544	465	501	585	375	346	160	601	300	0	726	487	241	655	
3	Bandarbar	316	437	0	469	391	427	510	301	271	85	526	225	121	651	413	167	581	
4	Rangamati	293	415	74	447	368	404	488	278	249	78	504	203	195	629	390	144	558	
5	Sajek	323	445	197	477	398	434	518	308	279	181	534	233	316	659	420	174	588	
6	Tanguar	163	305	349	346	281	277	173	132	217	319	279	188	425	248	212	240	154	
7	Nijhum Dw	376.8	440.8	98.1	433.8	369.7	333.8	562.7	348.9	341.3	130.2	580.7	278	34.7	708.2	490.7	219.4	639.6	
8	Saint Mart	470	591	200	623	544	580	664	454	425	239	680	379	79	805	566	320	734	
9	Dhuppani	257	271	31.7	238	214	184	411	218	201	65.2	388	172	117	523	295	124	436	
10	Ratargul	260.9	456.9	469		418.6	244	419.1	65.2	344.7	394.9	519.1	243.6	65.2	564.6	365.4	65.2	496	
11	Sundarban	242	98.7	314	99.3	146	171	323	299	208	271	191	265	297	412	195	256	377	
12	Lalakhal	291.4	487.4	499.5	513.2	477.4	441.5	449.6	203.5	375.2	425.4	549.6	274.1	574	595	395.8	337.6	526.4	
13	Austogram	370.1	566.1	578.3	620.3	556.2	491.9	425.9	282.2	425.6	504.2	525.8	352.8	652.8	571.3	474.6	416.4	502.7	
14	Potenga	259	380	102	413	334	370	454	244	215	17	470	168	177	595	356	110	524	
15	Guliyakhal	220	284	109	277	213	177	406	192	184	35	424	121	184	552	330	63	483	
16	Safari Park	48	299	288	349	272	232	174	114	166	246	274	147	347	320	151	204	251	
17	Sixty Dome	259	5.6	389	99	82	122	321	300	186	315	162	288	463	459	147	228	390	
18	Lawachara	192	388	400	442	350	314	351	104	276	326	450	175	475	496	297	239	427	
19	Sonargaon	32	218	305	271	207	171	206	91	105	231	230	86	379	351	136	143	283	
20	Zinda Park	17	281	321	287	254	187	187	93	121	247	287	102	395	332	134	159	263	
21	Moinot Gh	58	235	360	285	208	226	227	156	159	285	180	141	434	357	90	198	288	
22	Kuakata	349	183	412	54	111	145	469	342	225	338	349	271	486	607	237	251	539	
23	Mohamaya	190	254	139	247	183	147	376	162	155	65	476	91	214	522	300	40	453	
24	Chandrana	215	279	110	272	208	172	401	187	180	36	501	116	185	546	325	58	478	
25	Satchori N	140	336	349	391	304	291	299	52	224	274	399	123	423	444	245	187	375	
26	Kaptai	305	426	148	459	380	416	500	290	261	63	516	214	223	641	402	156	570	
27	Bhatiari	226	347	244	380	301	337	421	211	182	15	437	135	318	562	323	77	491	

Figure 5.1: Distance Dataset

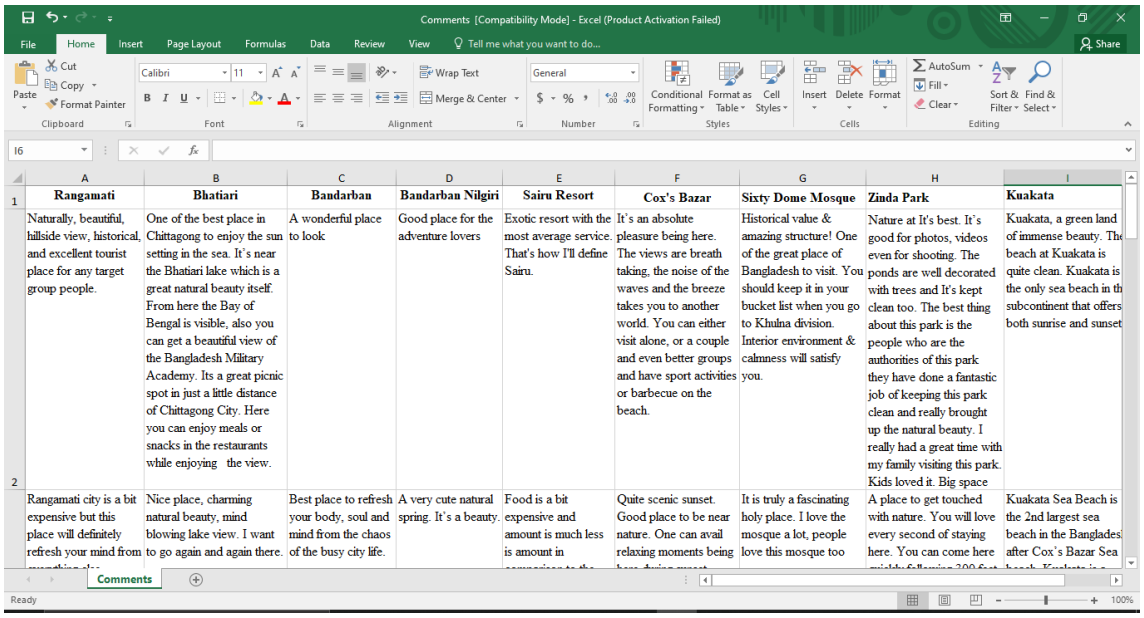
Firstly, for tour planning we have to know every tourism place's distance from every user who will search in our system. So, we cannot take the ideal distance from Dhaka to every place. We took every district in one dataset and as constant parameters, we selected all the tour places' names we gathered. The actual benefit of this dataset is, whenever any user will search about any place it will show the distance from his/her district. Like, if one user wants to visit Cox's Bazar, System will ask the user's district, Users replied with his/her location as Chittagong. It will show the distance of Chittagong to Cox's Bazar. Apart from this, it will show the fare of traveling from Chittagong to Cox's Bazar.
Secondly, we want to show some information to the user, so he/she can easily decide about his/her trip plan. In the second data set, we took all the tourist places of

Bangladesh and set the parameter as cost, Friends/family, season. For this season part, in our country, maximum tourist places are perfect for the winter and the rainy season.

1	Place	Cost	Friends=0/Family=1	Summer	Rainy	Winter
2	Keokradong	1200	1	0	1	1
3	Nafakhum	1500	0	0	1	1
4	Napittachora	1000	0	0	1	1
5	Sajek	2000	1	0	1	1
6	Damtua	1200	1	0	1	1
7	Tanguar	1000	1	0	1	0
8	Amiakhum	2000	0	0	1	1
9	Nijhum Dwip	1200	1	1	0	1
10	Saint Martin	1500	1	1	1	1
11	Shimul Garden	1000	1	1	1	1
12	Marayan Tang	1500	0	0	1	1
13	Hotongkuchu	1000	0	0	1	1
14	Dhuppani	2000	0	0	1	1
15	Ratargul	2000	1	0	1	0
16	Peyara Bazar	1000	1	0	1	0
17	Sundarban	3000	1	0	1	1
18	Lalakhal	1000	1	0	1	1
19	Austogram	1000	1	0	1	0
20	Potenga	500	1	1	1	1
21	Guliyakhali	800	1	1	1	1
22	Cox's Bazar	1100	1	1	1	1
23	Safari Park	500	1	1	1	1
24	Sixty Dome Mosque	500	1	1	1	1
25	Lawachara	1000	1	0	1	0
26	Sonargaon Tajmahal	500	1	1	1	1
27	Zinda Park	500	1	1	1	1
28	Moinot Ghat	500	1	1	1	1
29	Kuakata	1000	1	1	1	1

Figure 5.2: Dataset 2

For this, we divide it into three seasons as summer, rainy, and winter instead of showing results per month. Now, this dataset will show the minimum cost of that place per day. If a user wants to go Cox's Bazar from Mymensingh for three days- It will take a distance from dataset 1; which is 570km. Our travel cost per km is 1.42 taka. It will multiply this amount by distance to find travel costs. Travel cost is = 809 taka. Now, this user wants to stay there for three days, it will gather per day minimum medium level cost from dataset2. We set this price as medium level, so that all class people can understand this cost. Cox's Bazar per day cost is 800 for per person (hotel fare +two main meals+breakfast). Users will stay for three days. Three days cost is 800*3=2400. Now the total cost is =809+2400+809= 4016 taka for that person. The actual benefits of these two datasets are reliability and efficiency. We check this data from different travel groups. We collected this distance data from google map. We calculated this distance and cost from different places to check the accuracy and it works better than we thought.



	Rangamati	Bhatiari	Bandarban	Bandarban Nilgiri	Sairu Resort	Cox's Bazar	Sixty Dome Mosque	Zinda Park	Kuakata
2	Naturally, beautiful, hillside view, historical, and excellent tourist place for any target group people.	One of the best place in Chittagong to enjoy the sun setting in the sea. It's near the Bhatiari lake which is a great natural beauty itself. From here the Bay of Bengal is visible, also you can get a beautiful view of the Bangladesh Military Academy. Its a great picnic spot in just a little distance of Chittagong City. Here you can enjoy meals or snacks in the restaurants while enjoying the view.	A wonderful place to look	Good place for the adventure lovers	Exotic resort with the most average service. That's how I'll define Sairu.	It's an absolute pleasure being here. The views are breath taking, the noise of the waves and the breeze takes you to another world. You can either visit alone, or a couple and even better groups and have sport activities or barbecue on the beach.	Historical value & amazing structure! One of the great place of Bangladesh to visit. You should keep it in your bucket list when you go to Khulna division. Interior environment & calmness will satisfy you.	Nature at It's best. It's good for photos, videos even for shooting. The ponds are well decorated with trees and It's kept clean too. The best thing about this park is the people who are the authorities of this park they have done a fantastic job of keeping this park clean and really brought up the natural beauty. I really had a great time with my family visiting this park. Kids loved it. Big space	Kuakata, a green land of immense beauty. The beach at Kuakata is quite clean. Kuakata is the only sea beach in th subcontinent that offers both sunrise and sunset
	Rangamati city is a bit expensive but this place will definitely refresh your mind from	Nice place, charming natural beauty, mind blowing lake view. I want to go again and again there.	Best place to refresh your body, soul and mind from the chaos of the busy city life.	A very cute natural spring. It's a beauty.	Food is a bit expensive and amount is much less is amount in	Quite scenic sunset. Good place to be near nature. One can avail relaxing moments being	It is truly a fascinating holy place. I love the mosque a lot, people love this mosque too	A place to get touched with nature. You will love every second of staying here. You can come here	Kuakata Sea Beach is the 2nd largest sea beach in the Banglades after Cox's Bazar Sea

Figure 5.3: Comments Dataset

In this dataset, comments or reviews of general people about different places are gathered manually. These comments are being used for doing the process of opinion mining, emoji recognition and sentiment analysis. These processes are helped to sort the places which are getting by user's requirements. These comments are collected, uploaded and updated by admin manually.
We are updating this data on a regular basis so that it will be more efficient and reliable for showing data. We are thinking of adding more attributes here to show more replies to the user, Now The most important part is generating perfect output using our query.

5.2 Analytical Results

As our works mainly based on data science and data mining so our result is mainly based on how we are extract knowledge from our dataset. We used contextual filtering to fetch data from our dataset. In regular search the input is fixed and for that input you will get the output for queries. But in contextual filtering, we do not have to fix the input to get output. So, flexibility is there if we use contextual filtering. We used four types of filters in our algorithm which are filter based on trip time, filter based on budget, filter based on trip type and finally based on weather. Filter based on Trip Time and budget is mandatory as these 2 things are the main key of a trip. Trip type filter and weather filter is optional, user can choose whether to use it or not.
To filter using Trip day we have to consider the distance from users' destination to places that user can travel and as well as enjoy trip. So, we researched and come to conclusion which is shown below:

Place suggestion will be filtered with trip time given by customer based on the table shown above. It is totally a theoretical concept and value can be vary based con place, time and conditions.

Trip Duration	Travel Distance(KM)
1 Day	< 120
2 Day	<=175
3 Day	<=300
More than 3 Days	All Possible values

Table 5.1: Trip Time Condition Table

To filter using Budget we have to consider travel cost as well as cost of that place per day. We used per kilometer travel price as 4 taka/Km. And used distance database to calculate the trip cost. So, budget should be:

```
Journey cost = 2 * (Distance * Price Per KM)
    Place cost = Trip   Day * Cost Per Day
Total Cost for per Person = Journey Cost + Place cost
  Total Cost Per Person <= Budget Per Person
```

Then we applied filtering method based on trip type. Trip type is divided into two part which are friendly trip and family trip. Our dataset has a value of trip type which can distinguish between family trip and friendly trip. We applied the condition to filter places based on type of trip.

Last filter we used is weather filter. We divided season into 3 section which are Summer, Rainy and Winter. We have separate column in our database for weather which we used to filter places based on weather. If current season is winter, then our filter will filter data and will show only those places which are best for winter season. Similarly, system will show places which are best for summer season when current season is summer and will show places which are best for rainy season when current season is rainy. For some people season does not matter, they just want to travel and enjoy the trip. That's why we put this filter as optional, customer can choose to use it or can ignore it.

After we done with the filtering, we will sort and rank the filtered places based on sentimental analysis. So, in the next algorithm we used opinion mining algorithm, where we took, the reviews of each places from social media and applied sentimental analysis on it. We know social media is a great platform to get information about places they travel and can use it as a review. Then using that review we can give a rating based on sentimental analysis and sort the places according to sentimental result. So, user will get place suggestions in a sorted way where the first place will be the best and last will be worst based on social media review.

We divided our opinion mining into two parts. Comments in social media contents two things. Sentence which contains feeling and emojis which also contains emotions and feelings. Emoji is widely used nowadays and it has a significant value. So, we decided to use emoji in our sentimental analysis beside opinion mining.

To rate emojis, first we need to recognize emojis and separate them from sentence. As emojis do not have ASCII value, so we had to considered Unicode to recognize emojis

[11].At first, we read the comments or sentence from data-frame and separated emojis from sentence using Unicode. Then separated emojis where categorized based on positive reaction emojis and negative reaction emojis. Then we used this categorized emojis to do rating on emojis. We took 10% of the value of emojis rating, as we considered in a whole comments or reviews emojis carries 10 to 15 percent of values.

For opinion mining we had to do natural processing language. We already discussed about the six steps of NPL which are Tokenization, Stemming, Lemmatization, POS Tags, Named Entity Recognition, Chunking. There are a lot of works on NPL and for our opinion mining we do not have to use the whole process to do sentimental analysis. Rather we used textblob function of 'TextBlob Library' which will do all six process of NPL and will give us the sentimental result. TextBlob has all the function of NPL like Tokenization, Words Inflection and Lemmatization, Spelling Correction, Parsing and many more. Textblob function gives the result between -1 to 1. [32]

Rating Score	Review Type
Reviews < 0	Bad
Reviews > 0	Good
Review = 0	Neutral
Review = 1	Best
Review = -1	Worst

Table 5.2: Sentiment Rating Table

When we put the comments or reviews in the textblob function, it will process that sentence and will rate according to its conditions. So, we got the sentimental rating result of the filtered places using the comments and reviews of that places.

At last we merged the result of emoji rating and sentimental analysis of comments and reviews and sorted the places using this merged sentimental result. So, the place which has good reviews and comments will get highest sentimental score hence will be on top and the place which has negative reviews will get low score and will be in bottom. By this process user not only get the suitable places which fulfill his or her requirements but also finds the reviews of those places and can know which place is preferable by most of the people.

5.3 Operations of Emoji Recognition

Emojis were first developed as abstract depictions of emotions. Emojis are used, analogous to punctuation, to express sentiment, irony or laughter and to replace body language and voice tone in text-based communication. Emojis display emotions when they communicate feelings that make them significant. We also witness the conveyance of messages in our everyday lives. Emojis indicate that it is as necessary to provide emotional contact as using language. In addition to this, emoji are known as non-verbal resources. The emoji are the most effective communication instruments that enable individuals to communicate their linguistic ability.[10]

For this reason, we are considered emojis in all reviews and used emoji recognition process in order to identify emojis and try to find out their sentiment analysis. The steps of doing emoji recognition are given below:

- First of all, we need to separate emoji from a review because in the opinion mining process it cannot identify emojis rather it can identify sentences. That's why we need to separate emojis from reviews.
- Then we need to classify the emojis into two section which are- good and bad. It helps to understand the system which emoji contains what kind of meaning.
- Then we need to list all the emojis in an array which are collected from all comments or reviews.
- Then we need to compare the collected emojis with the classified emojis and have to do the rating process which will add with the result of sentiment analysis process.

Thus our whole process for emoji recognition will operate.

5.4 Operations of Opinion Mining

Opinion mining is key for both individuals and enterprises. Individuals may like to see other consumers' views on a product to evaluate it before purchasing it. In order to make future choices, businesses want to evaluate consumers' opinions on their goods. So, it is important to evaluate the viewpoint of the client and their reaction. In this way, enterprises will concentrate on enhancing their product characteristics that are not common with consumers. This contributes to overcome the criteria of marketing intelligence and product benchmarking in the manufacturing sector[2]
Considering these things, we decide to do opinion mining on collective comments or review for doing ranking process on the result of filtering places. The steps or operations of opinion mining are given below:

- We know that for doing opinion mining we have to deal with Natural Language Process (NLP). For doing opinion mining by using NLP we have to do tokenization, stemming, lemmatization, named entity recognition (NER) and chunking. But fortunately there is function in python which named as "TextBlob"; which has all the functionalities those are mentioned here. TextBlob makes our work easier.
- Firstly, the function "TextBlob" remove all capital letters, punctuations, emojis, links, etc. Essentially, to delete anything that is not words or numbers.
- Tokenize the data into terms, meaning that each comment is split into a set of individual words.
- Delete all stop words such as 'the', 'a', 'and', etc., which are words that do not add meaning to a statement.

- Remove all empty comments.
- It lemmatizes the text, which is the method of grouping the inflected forms of a word together so that they can be analyzed as a single object.
- Then do the POS operation to identify the parts of speech and to separate the adjective, adverbs which will use in the sentiment analysis process.

Thus the opinion mining process will be done and ready for the next process which is sentiment analysis.

5.5 Operations of Sentiment Analysis

The study of sentiment offers clues to what the most critical questions are. Since sentiment analysis can be automated, judgments can be taken based on a large amount of evidence rather than plain opinion that isn't always correct. Analysis the sentiments of opinion is important because firms tend to portray their brand favorably, or at least more positively, than competitors' brands. Using vast volumes of text data, sentiment analysis is useful for rapidly gathering insights.
After getting the results from opinion mining the sentiment analysis process will take place. The steps or operations of sentiment analysis are given below:

- By using textblob sentiment function, we can get the polarity of a sentence which contains that how positive or negative the sentence is.
- A float number between -1 and 1 which is supposed to be expressed whether a statement is positive (1) or negative (-1).
- The sentiment property of the textblob function is also able to return the subjectivity of a sentence. But in our project we will consider only the value of polarity in order to doing ranking of the places.

Thus the sentiments analysis process will complete and help us to rank the places based on the polarity of people's opinions or comments

Chapter 6

Implementation and Analysis

6.1 Algorithm

Our project is mainly based on contextual filtering and opinion mining algorithm. In first part we used contextual faltering algorithm to find the suitable place for user based on user input and then we used opinion mining algorithm on those places to measure sentimental result. Then sorted those places based on sentimental data.

We used panda library to read our data-frame and numpy library is used for array. Used re library for regular expression. Emoji library is used to do analysis on emoji and textblob library is used for sentimental analysis.

Libraries:

```
import pandas as pd
import numpy as np
import re
import emoji
from textblob import TextBlob
```

To filter based on trip time we used the table showed above and filled the condition in if-else statement to filter. In the algorithm, source is the location of user which is used to find the distance of source from destination.

Algorithm 1.1: Trip Time Filtering

Input Data → User_Location, User_TripTime

Read ← Dataset

if Data matches with Condition 1 **then**

    Execute Dataset → Filter dataset and select which has distance less than 120

    Save Place_Name ← Filtered places in Array List

else if Data matches with Condition 2 **then**

    Execute Dataset → Filter dataset and select which has distance less than 175

    Save Place_Name ← Filtered places in Array List

else if Data matches with Condition 3 **then**

    Execute Dataset → Filter dataset and select which has distance less than 300

    Save Place_Name ← Filtered places in Array List

else

    Execute Dataset → Select All the Places

    Save Place_Name ← Filtered places in Array List

To calculate cost, we used the formula:

```
Cost = (Trip Day * Cost Per Day) + 2 * (Distance * Price Per KM)
```

Based on this we done the filtering to find places where user can travel with his/her budget. Here 'dis' is new data set which we created based on the trip day condition. db1[Cost] is the data set cost of per day cost of tourist place and budget is the users' budget. Then applied the formula to check if any places are available under his/her budget or not. If our algorithm find places based on his/her budget then it will go to else condition and will show the places and will follow next process. If our algorithm does not find any places based on that budget then it will suggest to increase budget.

Algorithm 1.2: Budget Filtering

Input Data → User_Location, User_Budget

Execute Dataset → Select only Place Name and Cost from Dataset

Execute Cost → Calculate total Trip Cost

Execute Dataset → Filter dataset and select places which satisfy cost

if Dataset is not empty **then**

    Save Place_Name ← Filtered places in Array List

else

    Display ("No Place Found")

Then our algorithm will apply filter based on trip type. If user select family trip then if condition will be applied and will filter and select only those places which is suitable for family trip else will show the places where friend can go.

Algorithm 1.3: Trip Type Filtering

Input Data → User_TripType, Place_Name

if User_TripType matches with Family Trip Condition **then**

    for each values of Place_Name i **do**

        Execute Dataset → Filter Dataset based on Place_Name i

        Execute Dataset → Filter Dataset based on Family Trip Condition

        Save Place_Name ← Filtered places in Array List

else if User_TripType matches with Friendly Trip Condition **then**

    for each values of Place_Name i **do**

        Execute Dataset → Filter Dataset based on Place_Name i

        Execute Dataset → Filter Dataset based on Friendly Trip Condition

        Save Place_Name ← Filtered places in Array List

else

    Save Place_Name ← Filtered places in Array List

Next filtering will be done based on weather. But this filter is optional and will only do if user select the option top filter based on weather. We divided weather into 3 section which are summer, rainy and winter. If user choose the option of weather filtering then algorithm will look for the weather. If it is summer then it will run the if condition weather = summer, if its rainy then it will run else if condition weather = rainy and if it is winter it will run next else if condition weather = winter. Based on this condition it will filter the result.

Algorithm 1.4: Seasonal Filtering

Input Data → Place_Name, Season

if Season matches with Summer **then**

    for each values of Place_Name i **do**

        Execute Dataset → Filter Dataset based on Place_Name i

        Execute Dataset → Filter Dataset based on Summer Condition

        Save Place_Name ← Filtered places in Array List

else if Season matches with Rainy **then**

    for each values of Place_Name i **do**

        Execute Dataset → Filter Dataset based on Place_Name i

        Execute Dataset → Filter Dataset based on Rainy Condition

        Save Place_Name ← Filtered places in Array List

if Season matches with Winter **then**

    for each values of Place_Name i **do**

        Execute Dataset → Filter Dataset based on Place_Name i

        Execute Dataset → Filter Dataset based on Winter Condition

        Save Place_Name ← Filtered places in Array List

else

    Save Place_Name ← Filtered places in Array List

We will do opinion mining on comments of those places which is generated in previous algorithm. At first, we will do emoji rating. To rate the emoji, we will separate emojis from comments. Here we classified 2 section which are good emoji and bad emoji. We used regular expression to separate emojis. Then we listed all the

emojis of that comment in an array. Then we compared it with our 2 section of emojis and did the rating. We took only 10% values of emoji to use it with sentimental analysis.

Algorithm 2.1: Emoji Detection

Input Data → Comments or Reviews

function demoji (Data)

    Read emoji_pattern → All Pattern of Emoji

    Execute Data → Match Data with Emoji Pattern

    Save Emoji ← Save Emoji in List

    Return Emoji

end function

Algorithm 2.2: Emoji Extraction

Input Data → Comments_Reviews, Place_Name

for each values of Place_Name i **do**

    Execute Dataset → Filter Dataset based on Place_Name i

    for each Comments_Reviews j **do**

        Execute text → Detach emojis using demoji function for Comments_Reviews j

        Execute text → findall regular expression function to remove unnecessary sign

        Execute list_emoji → Use emojize function to separate and create emoji list

        Save list_emoji ← Filtered places in Array List

Algorithm 2.3: Emoji Rating

Input Data → list_emoji

Read Good_Emoji → All Positive Emojis

Read Bad_Emoji → All Negative Emojis

for x in list_emoji **do**

    if x matches with Goodemoji **then**

        Execute emo_rating → Increase Values of Rating

    else if x matches with Bademoji **then**

        Execute emo_rating → decrease Values of Rating

    else

        Execute emo_rating → Do nothing

Execute Avg_Rate → Calculate average rating value of emojis

Save Emoji_Rate ← Take 10% Values from Emojis Rate

Then we used textblob function to find sentimental analysis on each comment. For each comment textblob function will give a value from -1 to 1 where less than 0 is bad review and greater than 0 is good review. Using this result we find the sentimental result and added it with emoji rating to get an overall score for that place. Higher the score is better the review will be of that place. This is the part where opinion mining has been done.

Algorithm 3: Opinion Mining

Input Data → Place_Name, Comments_Reviews

for each values of Place_Name i **do**

    Execute Dataset → Filter Dataset based on Place_Name i

    for each Comments_Reviews j **do**

        Execute Text → Put each Comments and Reviews in TextBlob Function

        Execute Text_Pol → Calculate Polrity of Each Comments and Reviews

        Execute Sentimental_result → Calculate total sentimental result of each place

        Execute Result → Merge Emoji Rating and Sentimental Rating

        Save Total_Score ← Store total score of each place

Sentimental Analysis:

Algorithm 4: Sorting Based on Sentimental Analysis

```
Input Data → Total_Score, Place_Name
for each values of Place_Name i do
        for each values of Place_Name j do
                if value of Total_Score i is greater than Total_Score j then
                        Swap Total_Score i with Total_Score j
                        Swap Place_Name i with Place_Name j
Display (Place_Name)
```

At last our algorithm will sort the places based on sentimental analysis value. This part will only do sorting. Those places with higher sentimental value will go first/up and those will less sentimental value will go next/last.

At last sorted result will be printed in a text file which we will use in our website/app as a result.

6.2 Implementation

Here we have created a simple layout of our website which contains some of the main features of the website. We have created this from wix.com. At first, for the homepage we have chosen a blank template and then used a header title to create the name of our website name. Then we have used a black strip to prepare slideshow of the places to visit.

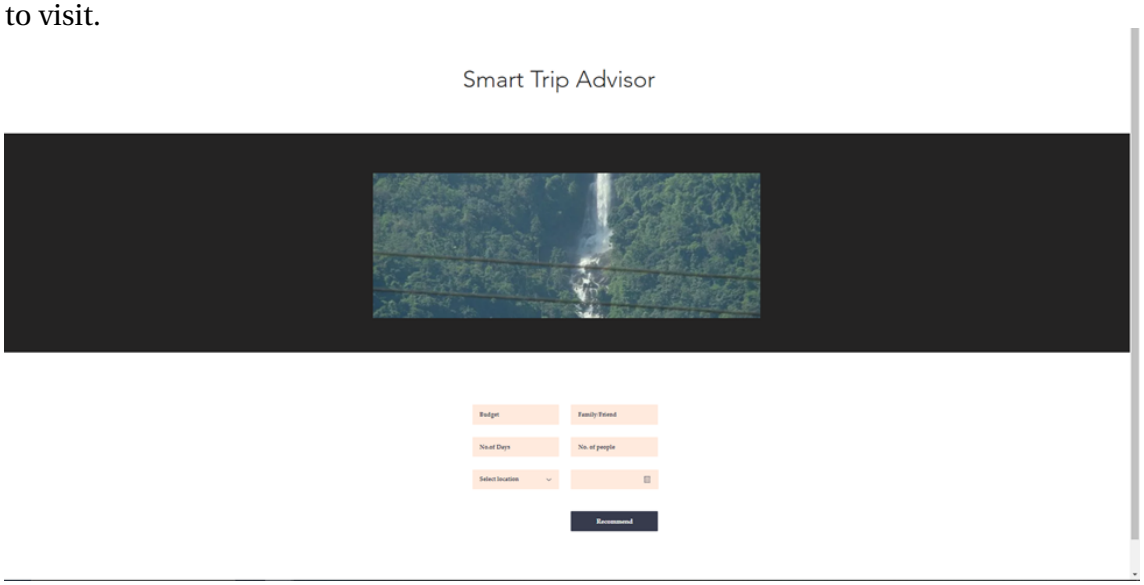


Figure 6.1: Home Page

After that, comes one of the main features, below the strip there is a form, which is one of the main important feature of the website. This form contains budget, no. of days, family/friend, type of location and finally the season. Here the basic contextual filtering happens for sorting out the location. For example, if the preferred season is about summer, then it consider the given condition and matches it with the other filters then it would recommend about those place.

In the services page we then rank the selected places with sentimental analysis of the comments about the place that we collected. To build the page we used to a preset template from the services tab on wix.com and in Fig2. It can be seen that the priority of the places would display from left to right. Along with this it also give a small summary of the place. In addition to recommending the places, we have also created a text field to give a brief description about the place. So that the users might be able to get a general idea about the places.

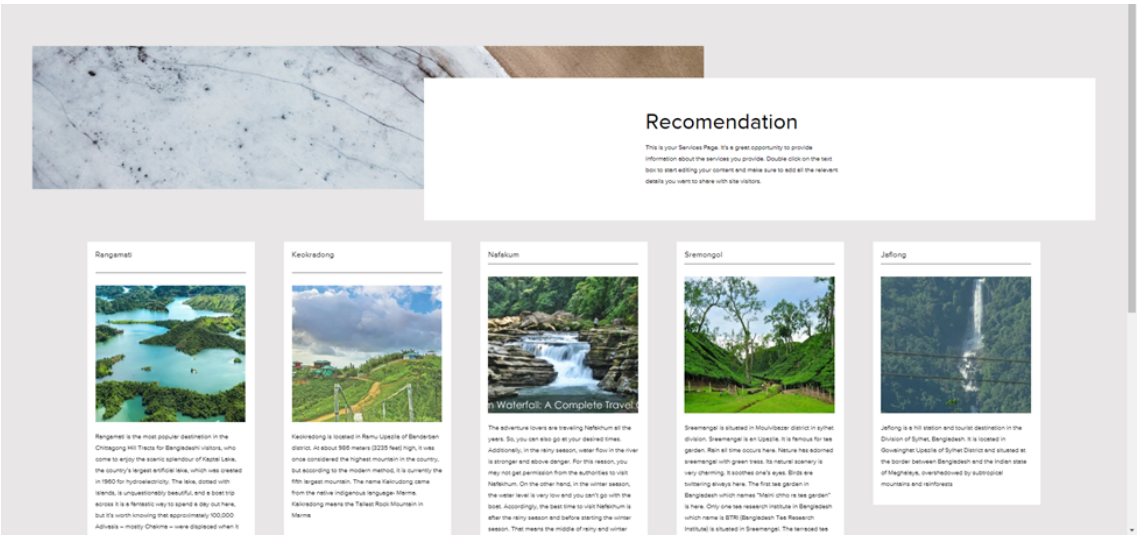


Figure 6.2: Recommended places

Here we have showed the output which is getting by the filtering algorithm. Our filtering algorithm will give us for this input. Based on this inputs our algorithm will apply filters on dataset and will generate the place name based on the input.

After getting the result of contextual filtering. Our process will apply another algorithms which are emoji recognition, opinion mining and sentiment analysis for doing the ranking process on the above mentioned places. Our algorithm will separate sentence and emojis, then it will operate to find sentimental result of reviews as well as rating of emojis. Below we showed couple of outputs for inputs of reviews in table. Table represent the output our algorithm will give in background to process our system. This output will be computed in the background and based on this our other process will be done.

In the Table 6.2, we gave the comments with emojis in the input of our algorithm and our algorithm separated the emojis. After that our algorithm will rate them and give both emojis and rating of emojis as output.

Inputs	Outputs
Location: Chittagong Budget : 2000 Trip Duration : 1 Trip Type : Both Weather : None	Rangamati Potenga Guliyakhali Mohamaya Lake Chandranath Kaptai Bhatiari
Location : Dhaka Budget : 3000 Trip Duration : 2 Trip Type : Friendly Season : Winter	Safari Park Sonargaon Tajmahal Zinda Park Moinot Ghat
Location : Khulna Budget : 1500 Trip Duration : 1 Trip Type : Both Season : Summer	Sixty Dome Mosque
Location : Barisal Budget : 3500 Trip Duration : 3 Trip Type : Both Season : Rainy	Sixty Dome Mosque Sonargaon Tajmahal Moinot Ghat Kuakata Mohamaya Lake

Table 6.1: Contextual Filtering

Comments	Emojis	Ratings
Old bridge, Narrow road, Risky😰😰. Crowded and so many people around there. 😪😫	'😰', '😰', '😪', '😫'	-0.12511
This lake is a beauty to watch❤️. If you roam over the lake by boat 😻 you will definitely want to revisit it. 👌	'❤️', '😻', '👌'	0.112
I went there as I am die-hard fan 🙂. But really disappointed to see the place 😣. My expectation was more than that 😕. Entry fee seems little higher 👎	'🙂', '😣', '😕', '👎'	-0.1111
Transportation is very poor 😩 but overall nice place just love it 😍.. ♡	'😩', '😍', '❤️'	0.109
The Temple was ordinary like others to look at 😕. I liked much the whole way 😊! The way was very beautiful and attractive with so many hills, trees and stairs😊! But the final view from top of the hill was best🤗😊.	'😕', '😊', '❤️', '🤗', '😊'	0.111090

Figure 6.3: Emoji Rating

In this Table 6.3, after getting the ratings for emojis; we have put the comments in the input of our algorithm and our algorithm gave the output showed in sentimental result and subjectivity result. For each comments the output is generated between -1 to 1 based on comments type.

Later, after merging these two values of sentimental result and emoji rating, we sorted the places based on these values.

Comments	Sentimental Results	Subjectivity Results	
In Bangladesh there are very few hills. In Hill place Keokradong is much known natural beauty place. People who love Hill tracking they enjoy the place and its really very good place for tracker. Thanks.	0.2928571429	0.4014285714	
One of the highest peaks of our Country-Bangladesh. But the roads are risky and not smooth. They are full of dust. You can't get to the peak with vehicle in the season of Rain. But you can walk and climb the peak. It is a place of great adventure! If you visit the place, you should visit in a moonlit night.	0.3833333333	0.6	
The biggest waterfall of Bangladesh. People can't express the beauty of Damtua in words, they just can see it amp; feel it inside!!! Though, it's really very hard to go there because of its remote location. . . ..If someone's health is not fit enough to hike 2-3 hours on Hilly trail, then I will suggest them not to go there...!!!	-0.1760416667	0.4510416667	
Outstanding natural view indeed. One of the best places I have visited till now Place is great, view is awesome and environment is excellent. Felt like heaven!!!!	0.73333333	0.7208333333	
It's a different kind of beauty. Hills are surrounded with water and there are lot of small island. The drone view would be mesmerizing. Shuvolong-1 and Shuvolong-2 waterfalls are main attractions there. height	0.1633333333	0.5866666666	

Table 6.2: Sentiment Result

6.3 Feature Analysis

In our research our target is providing people the best place to travel using their requirement. Here many people face different kinds of problems in their real life events. So we took a very short data collection from general people to identify what they prefer most and what they faced in their previous tours. To do so we make a questionnaire of over 15 attributes where the participants have to answer in a binary form. It helps us to identify their opinion. We asked the same question to different general people like students or job holders or businessmen to find general answers from them. Here are some of our important questions and their responses:

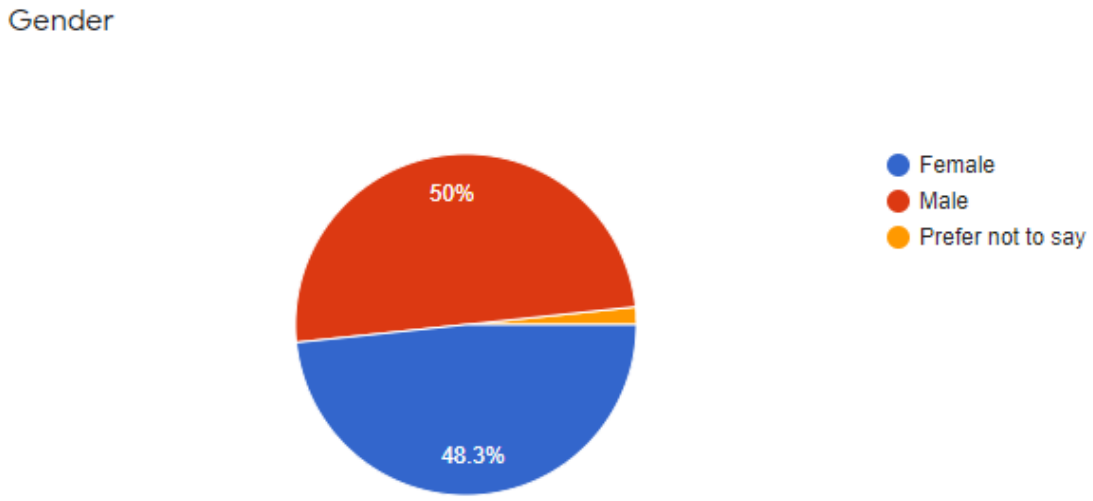


Figure 6.4: pie chart response of gender collected from general people

So in our targeted people, the numbers of both genders are almost equal thus we can justify our best answers to both genders.

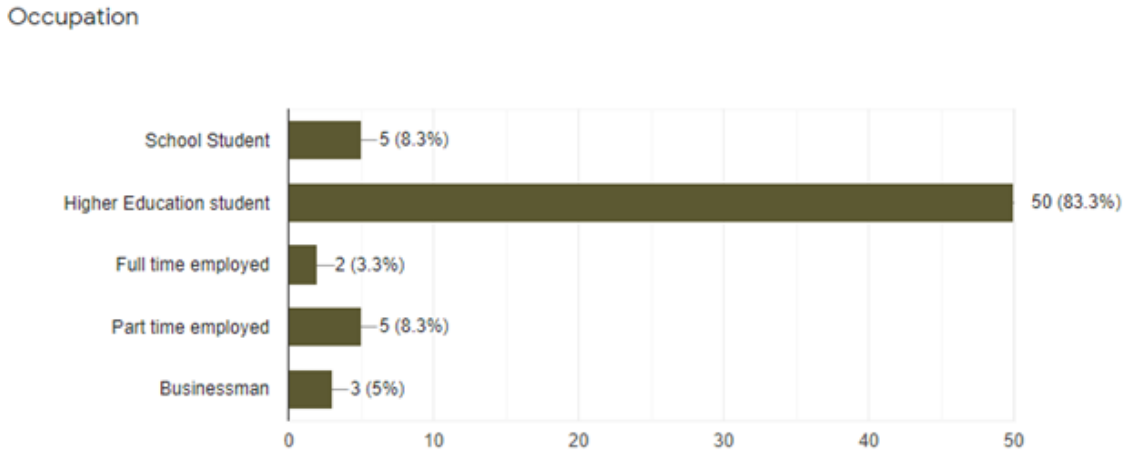


Figure 6.5: Bar response of Occupation collected from general people

Here, in our responses we have all kinds of people but the maximum number of

responses are from higher education students. But we collected data from employers and businessmen also.

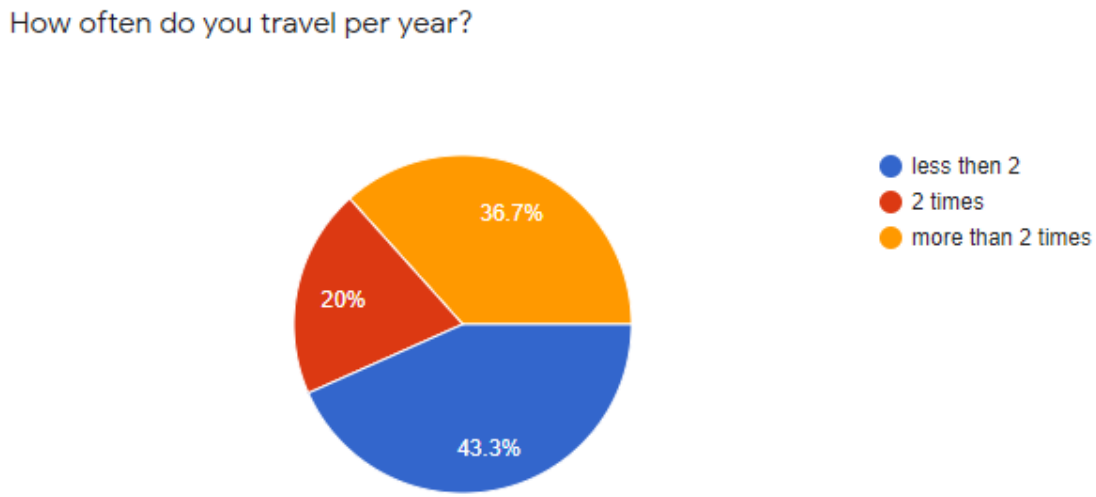


Figure 6.6: Pie chart response of peoples travelling rate per year

Here we can see that, maximum people, 43.3% people travel less than 2 times per year. Thus for they explore little less than others. So they need the proper guidelines so that they can enjoy their time. Any sort of misguide can destroy their time and money. So before traveling, they need proper suggestions, proper budget and proper time schedule so that they can cover everything in time.

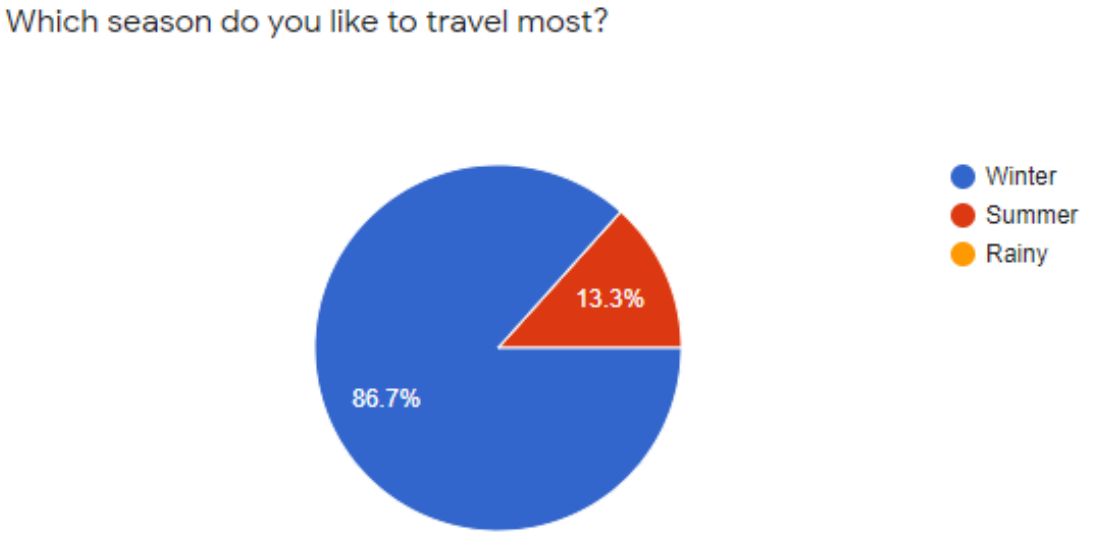


Figure 6.7: Pie chart response of general people preferences of travelling season

In our country, basically people like to travel in winter most. But this is the wrong thing people do always. That's why they need some suggestions, Because Many people don't know which season is best for which place. Normally in our country there are

lots of waterfalls. In the rainy season they get their best beauty. In winter it loses water and beauty. So the best time is the rainy season to enjoy the beauty of waterfalls. But none of our participants voted for the rainy season. That's why they need proper suggestions before they go on a trip.

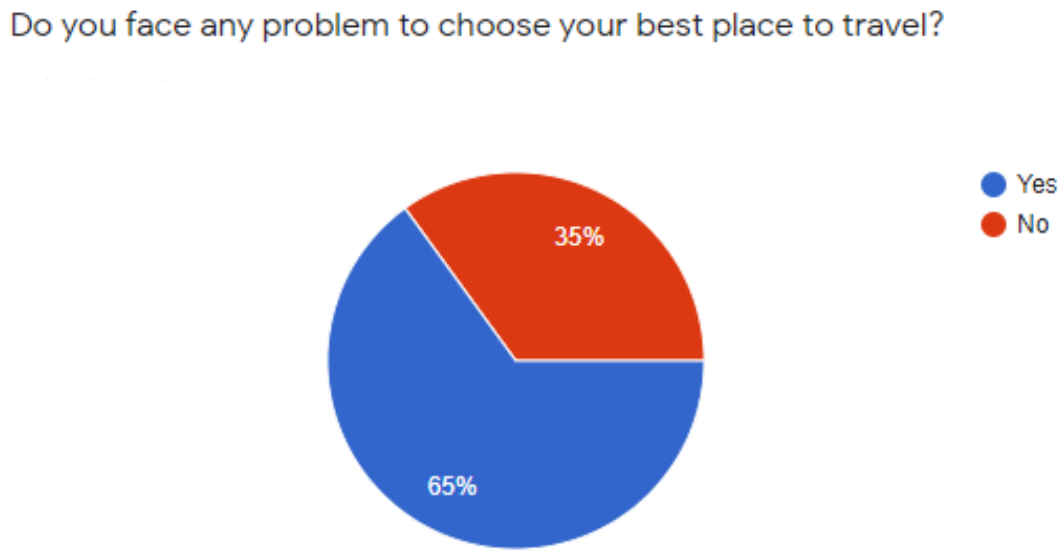


Figure 6.8: Pie chart response of general people preferences of travelling season

Here, 65% of people are facing different problems to choose their best place. People want to travel but they don't have enough information to travel perfectly. So they can't travel to the desired places. Often they waste their valuable time and money to enjoy but they end up with frustration and bad experience. So we need a better solution to help them to choose the best place.

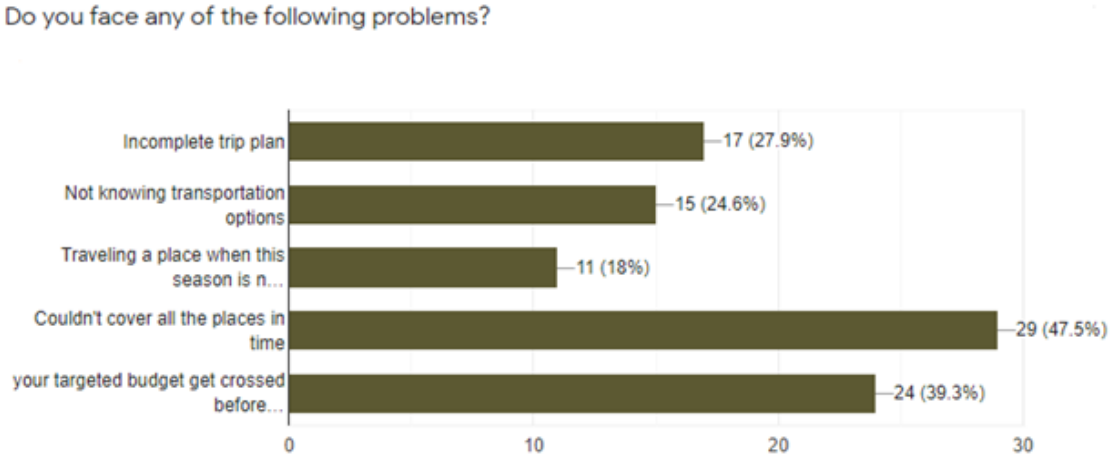


Figure 6.9: Bar diagram responses of following problem

In these questionnaires we set five different common problems which have been faced by many tourists in our country. The first one is an incomplete trip plan. This is a common issue in our country. People go on a trip without any plan and

thus they miss a lot of chances to enjoy. In our responses almost 28% of people voted for this one. The second one is not knowing transportations options, for this people face a lot of problems and the local peoples take double money from them. Which is ultimately a bad movement of a tourist. Almost 25% of people vote for this option. The third option is traveling to a place at the wrong time. People think that winter is perfect. But winter is not perfect for all the places. 18% of people voted for this. The fourth option is couldn't cover all the places. This is a very common problem. They don't have a plan and they miss lots of places. The highest number of people voted for this problem which is almost 48%. The last option is to get crossed. When people go for a budgeted trip and they need to use their money wisely. Many people don't have the right information about money, so they fail to complete their trip perfectly. Almost 39% of people voted for this option. Now we want to provide people exact information about transportation and cost of every single thing and provide them a best trip plan so that they can cover all the places they want to visit.

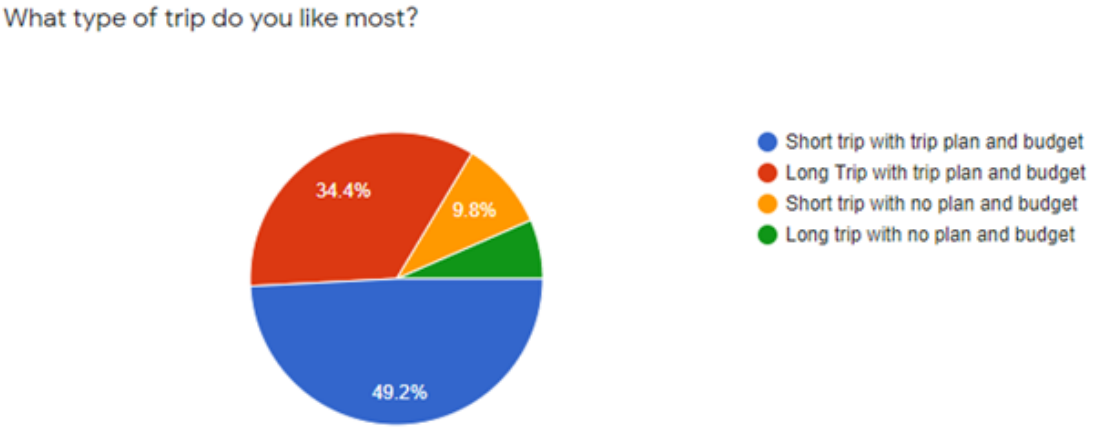


Figure 6.10: Pie chart response of trip which people like most

So, in result people like most Short trips with a plan and budget, almost 49.2% people like this. A long trip with a budget and plan is 34.5%. Without a plan for a short and long tour both percentages are below 10.

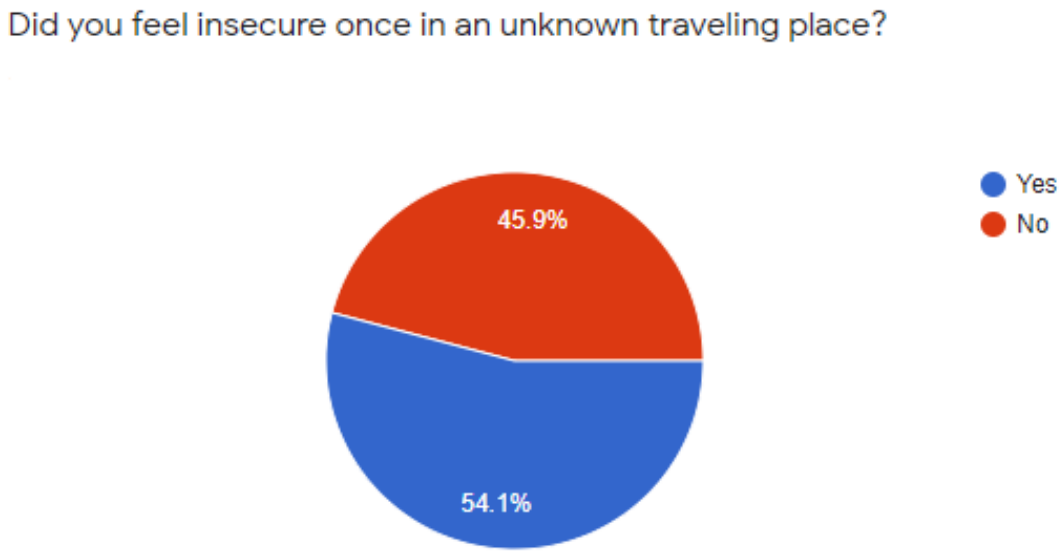


Figure 6.11: Pie chart response of feeling insecure

Here, 54.1% feel insecure in an unknown travelling place. The first thing a tourist needs is security. We have to ensure security to every tourist.

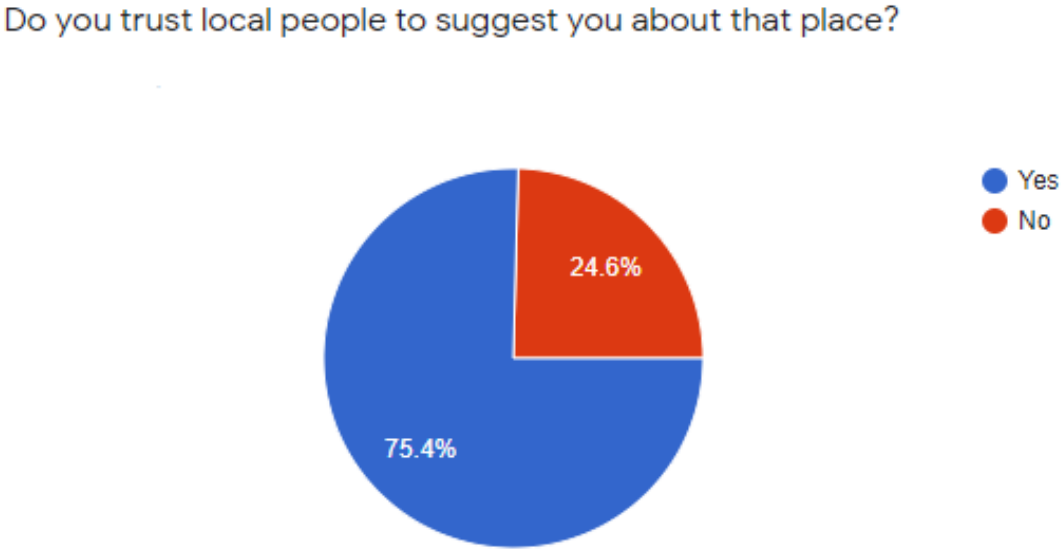


Figure 6.12: Pie chart response of local people are trustworthy or not

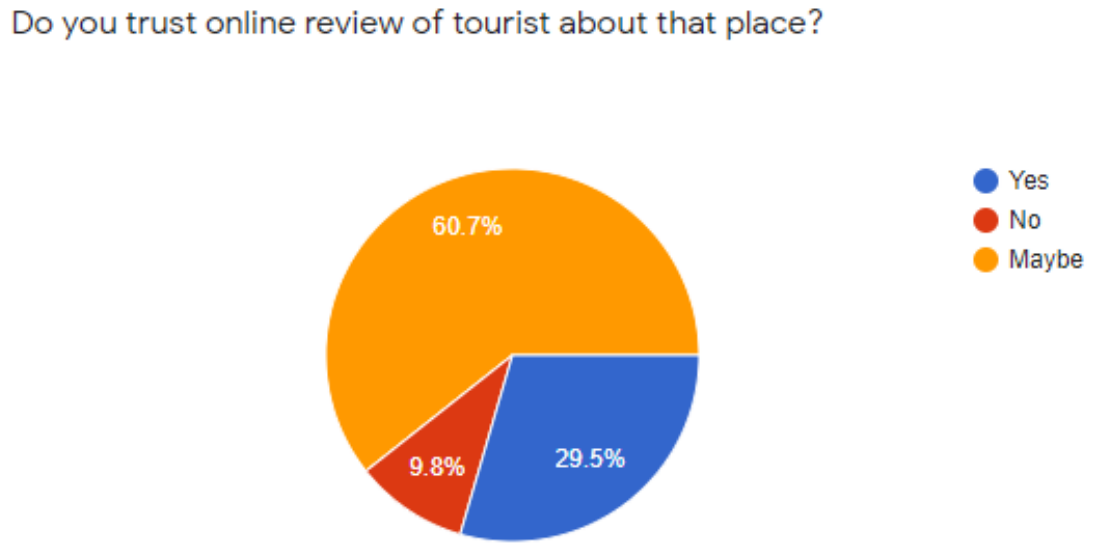


Figure 6.13: Pie chart response of online reviews are trustworthy or not

We got the lowest percentage at answer no. Maximum people voted for may be that means they don't trust them but they got the idea or plan from that review.

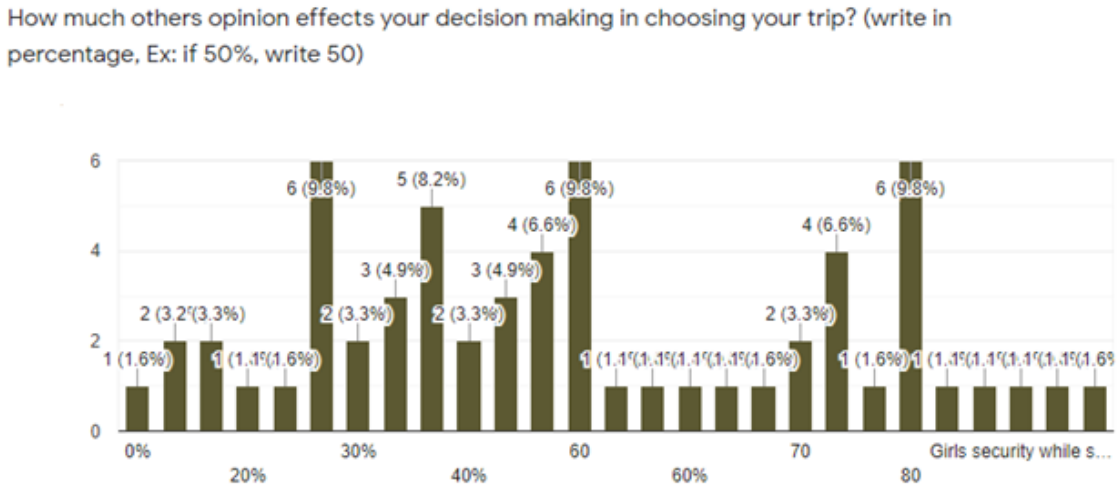


Figure 6.14: Bar diagram responses of others opinion effects

Here maximum people responded above 50%. A few people responded 0-30%. There are some people in between 40-50%.

Figure 6.15: Pie chart responses of their expected trust on an online platform to suggest trip

Here 83.6% people want to get help about their trip plan and suggestion. It will really help them a lot to enjoy their time with budgeted money. Therefore we want to build a system which will help them to suggest the best place.

Chapter 7

Conclusion and Future Research

7.1 Conclusion

In our research, we have designed, implemented as well as introduced a system where user can find out the most suitable places for having a great trip according to their requirements. Along with this, users are able to know that which place is more suitable for their trip which we are presented based on people's opinion. We have used the contextual filtering algorithm to have preferable places according to the user's requirements. We have also used the emoji recognition process, opinion mining and sentiment analysis process in order to analyze people's opinion towards a place. Based on those analysis we have done a sorting process to rank the places where user can able to know that which place is more suitable for a trip. We have used the knowledge of Data Science, Machine Learning and Natural Language Process to complete our system.

7.2 Limitations

Despite the fact that we thought about all significant potential standards traveling, still there are a few territories we didn't or couldn't cover. Traffic is one of them that we needed to cover yet proved unable. We needed to show the traffic state of the street on outing day but since of absence of data set and furthermore on account of this pandemic circumstance of the world we were unable to gather legitimate data set to cover this part. Additionally, we didn't consider the occasional value change of outing compels as it will make the cycle more troublesome. We considered major compels however there are some obligations we didn't consider or couldn't consider which we will attempt to include in our framework in future.

7.3 Future Work

In the future, we need to make our framework an AI-based framework where our framework can give the consequence of the inquiry with its scholarly information without utilizing the data set. We are intending to utilize Machine Learning to prepare our framework so it can respond and give arrangements all alone without utilizing

an information base. Fundamentally, what we planned to do is, we will prepare our machine utilizing the data set produced by the client's information and yield our framework. At the point when our framework will create enough information to prepare our machine then our machine can settle on choice all alone without the assistance of the data set. Additionally, we will keep both cycles parallel so the framework can learn just as well as improve results step by step. Additionally, we are intending to improve our separating condition to improve results.

References

[1] L. Lee, *Opinion mining and sentiment analysis*, 2008. [Online]. Available: https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf.

[2] B. Seerat and F. Azam, "Opinion mining: Issues and challenges (a survey)," *International Journal of Computer Applications*, vol. 49, pp. 42–51, Aug. 2012. DOI: 10.5120/7658-0762.

[3] P. S and S. F. S, "Opinion mining and sentiment analysis - an assessment of peoples' belief: A survey," *International Journal of Ad hoc, Sensor Ubiquitous Computing*, vol. 4, no. 1, pp. 21–33, 2013. DOI: 10.5121/ijasuc.2013.4102. [Online]. Available: https://cutt.ly/1jf5WFM.

[4] K. Khan, B. Baharudin, A. Khan, and A. Ullah, "Mining opinion components from unstructured reviews: A review," *Journal of King Saud University - Computer and Information Sciences*, vol. 26, no. 3, pp. 258–275, 2014, ISSN: 1319-1578. DOI: https://doi.org/10.1016/j.jksuci.2014.03.009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S131915781400010X.

[5] H.-S. Chiang and T.-C. Huang, "User-adapted travel planning system for personalized schedule recommendation," *Information Fusion*, vol. 21, pp. 3–17, 2015. DOI: 10.1016/j.inffus.2013.05.011.

[6] O. P. Patri, K. Singh, P. Szekely, A. V. Panangadan, and V. K. Prasanna, "Personalized trip planning by integrating multimodal user-generated content," in *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015)*, 2015, pp. 381–388. DOI: 10.1109/ICOSC.2015.7050837.

[7] W. Y. Kwon, Mingu Kim, and I. H. Suh, "Probabilistic tourist trip-planning with time-dependent human and environmental factors," in *2016 International Conference on Big Data and Smart Computing (BigComp)*, 2016, pp. 505–508. DOI: 10.1109/BIGCOMP.2016.7425980.

[8] X. Li, M. Li, Y. Gong, X. Zhang, and J. Yin, "T-desp: Destination prediction based on big trajectory data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2344–2354, 2016. DOI: 10.1109/TITS.2016.2518685.

[9] L. Ravi and S. Vairavasundaram, "A collaborative location based travel recommendation system through enhanced rating prediction for the group of users," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–28, 2016. DOI: 10.1155/2016/1291358.

[10] MohammedRajhi, *Emotional recognition using facial expression by emoji in real time*, 2017. [Online]. Available: https://www.grin.com/document/379502.

[11] 2018. [Online]. Available: https://stackoverflow.com/questions/43146528/how-to-extract-all-the-emojis-from-text.

[12] S. Migliorini, D. Carra, and A. Belussi, “Adaptive trip recommendation system: Balancing travelers among pois with mapreduce,” in *2018 IEEE International Congress on Big Data (BigData Congress)*, 2018, pp. 255–259. DOI: 10.1109/BigDataCongress.2018.00045.

[13] M. Preisendorfer, S. A. Sengupta, J. A. White, and A. A. Tekeoglu, “Social media emoji analysis, correlations and trust modeling,” Twitter is an ever-growing social-media platform where users post tweets, or small messages, for all of their followers to see and react to. This is old news of course, as the platform first launched over ten years ago. Currently, Twitter handles approximately six thousand new tweets every second, so there is plenty of data to be analyzed. With a character limit of 140 per tweet, emojis are commonly used to express feelings in a tweet without using extra characters that more explainingmight use. This is helpful in identifying the mood or state of mind that a person may have been in when writing their tweet. From a computing standpoint, this makes mood analysis much easier. Rather than analyzing a group of words and predicting moods from keywords, we can analyze single (or many) emoji(s), and then match those emojis to commonly expressed emotions and feelings. The objective of this research is to gather large amounts of Twitter data and analyze emojis used to find correlations in societal interactions, and how current events may drive social media interactions and behaviors. By creating topic models for each user and comparing it with the emoji distribution analysis, a trust ”fingerprint” can becreated to measure authenticity or genuineness of a given user and/or group of users. The emoji distribution analysis also provides the possibility of demographic predictions. Analysis is not limited to Twitter of course but is used here becausethe API is free and generally easy to use. This paper aims to prove the validity of emoji analysis as a method of user identification and how their trust models can be used in conjunction with pre-existing models to improve success rates of these models., Ph.D. dissertation, 2018. [Online]. Available: http://hdl.handle.net/20.500.12648/1090.

[14] J. Wang, X. Kong, W. Zhao, A. Tolba, Z. Al-Makhadmeh, and F. Xia, “Stloyal: A spatio-temporal loyalty-based model for subway passenger flow prediction,” *IEEE Access*, vol. 6, pp. 47 461–47 471, 2018. DOI: 10.1109/ACCESS.2018.2865921.

[15] 2019. [Online]. Available: https://www.dataversity.net/a-brief-history-of-machine-learning/#.

[16] 2019. [Online]. Available: https://app.amanote.com/v3.9.15/note-taking/document/DotV0nMBKQvf0Bhi6Q5N.

[17] 2021. [Online]. Available: https://en.wikipedia.org/wiki/Machine_learning.

[18] 2021. [Online]. Available: https://en.wikipedia.org/wiki/Data_science#:~:text=The%5C%20term%5C%20%5C%E2%5C%80%5C%9Cdata%5C%20science%5C%E2%5C%80%5C%9D%5C%20has,definition%5C%20was%5C%20still%5C%20in%5C%20flux.

[19] 2021. [Online]. Available: https://www.dataversity.net/brief-history-data-science/.

[20] 2021. [Online]. Available: https://www.oreilly.com/library/view/doing-data-science/9781449363871/ch01.html.

[21] 2021. [Online]. Available: https://en.wikipedia.org/wiki/Natural_language_processing.

[22] 2021. [Online]. Available: https://en.wikipedia.org/wiki/History_of_natural_language_processing.

[23] 2021. [Online]. Available: https://nlp-now.co.uk/nlp-history/?doing_wp_cron=1610100965.8982949256896972656250#:~:text=NLP%5C%20grew%5C%20out%5C%20of%5C%20the,California%5C%2C%5C%20Santa%5C%20Cruz%5C%20in%5C%201970.

[24] 2021. [Online]. Available: https://www.landsiedel.com/en/nlp/history-of-nlp.html#beginning.

[25] 2021. [Online]. Available: https://www.hindawi.com/journals/cin/2016/1281379/.

[26] 2021. [Online]. Available: https://en.wikipedia.org/wiki/Emoji.

[27] B. Marr, *Ashorthistoryofmachinelearningeverymanagershouldread*, 2021. [Online]. Available: https://www.forbes.com/sites/bernardmarr/2016/02/19/a-short-history-of-machine-learning-every-manager-should-read/?sh=42e2523d15e7.

[28] G. Press, *A very short history of data science*, 2021. [Online]. Available: https://www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/?sh=57c0dc5155cf.

[29] T. Start? And V. Gladchuk, *A history of machine learning from 1940s to present days*. 2021. [Online]. Available: https://labelyourdata.com/articles/history-of-machine-learning-how-did-it-all-start/.

[30] U. Team, *A modern history of data science*, 2021. [Online]. Available: https://datasciencedegree.wisconsin.edu/blog/history-of-data-science/.

[31] C. William S., *William s. cleveland*, 2021. [Online]. Available: https://www.stat.purdue.edu/~wsc/.

[32] [Online]. Available: https://textblob.readthedocs.io/en/dev/quickstart.html.

Chapter 8

Appendix A: Prepared Questionnaires

The questionnaires which are being used for data collections are given below:

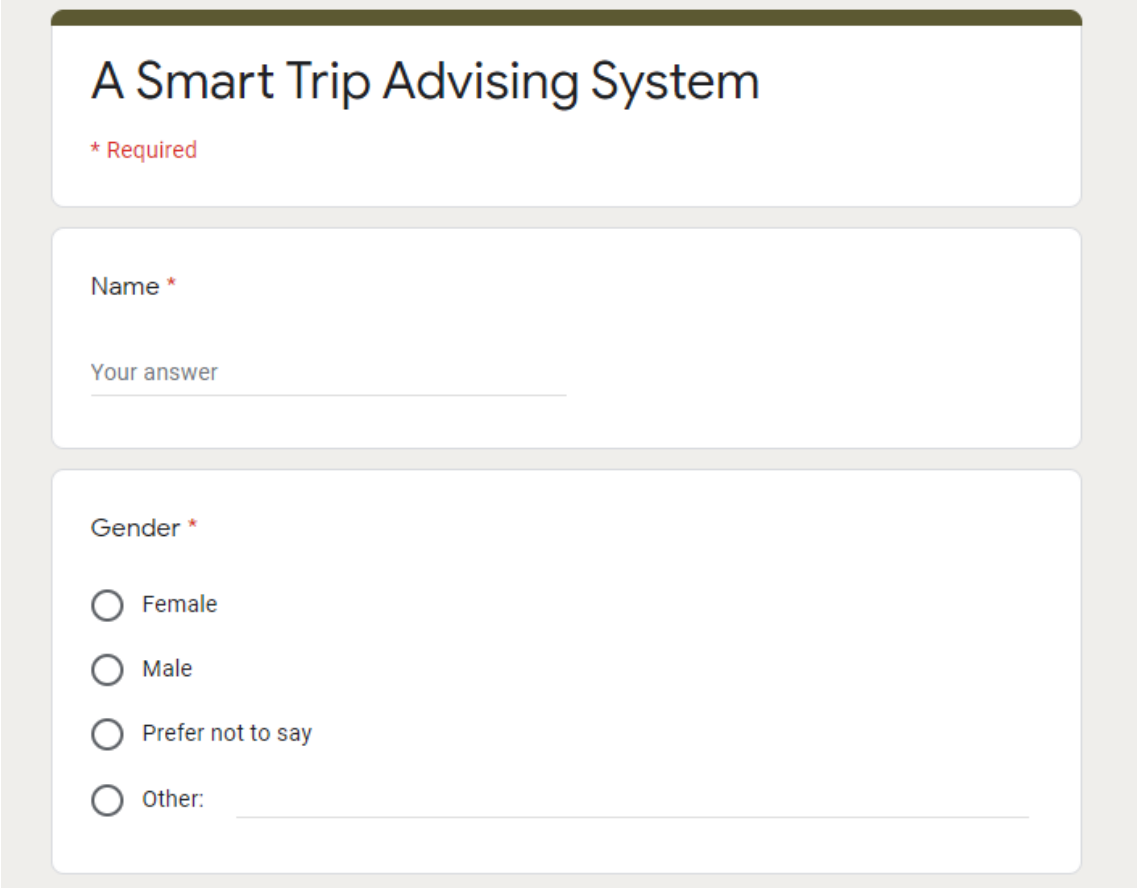


Figure 8.1: Part 1

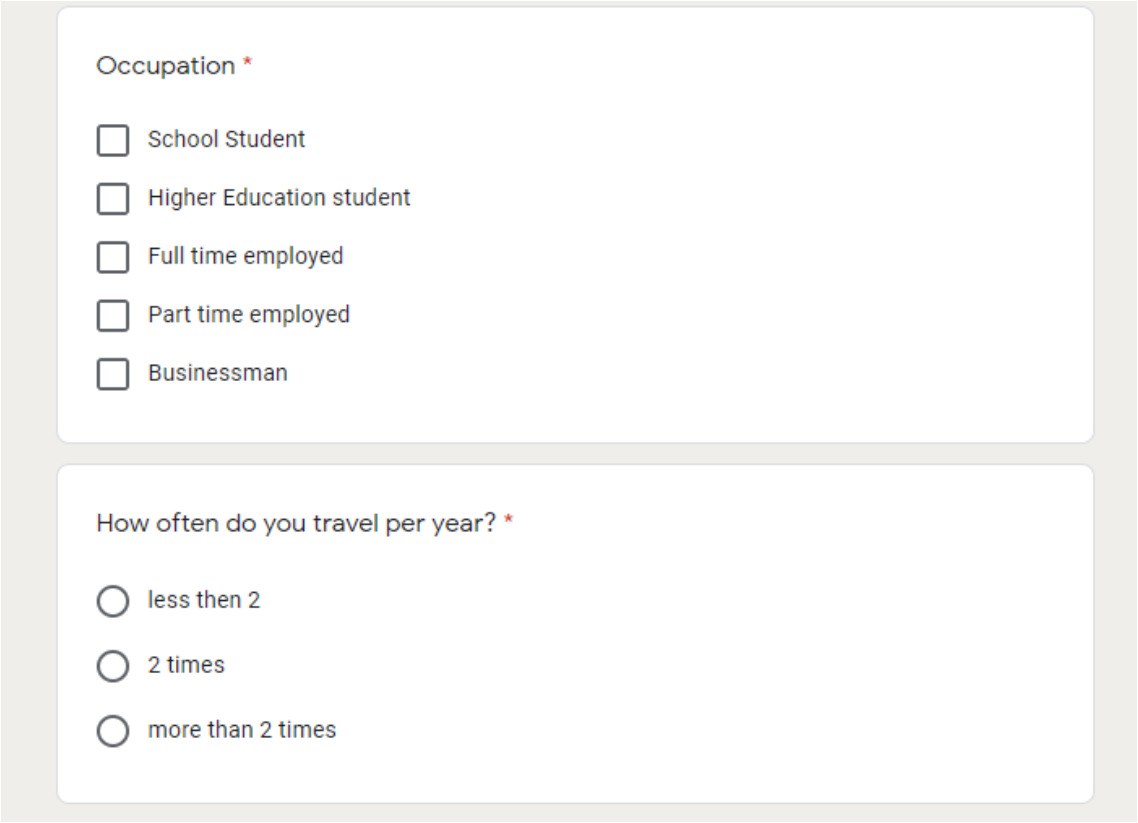


Figure 8.2: Part 2

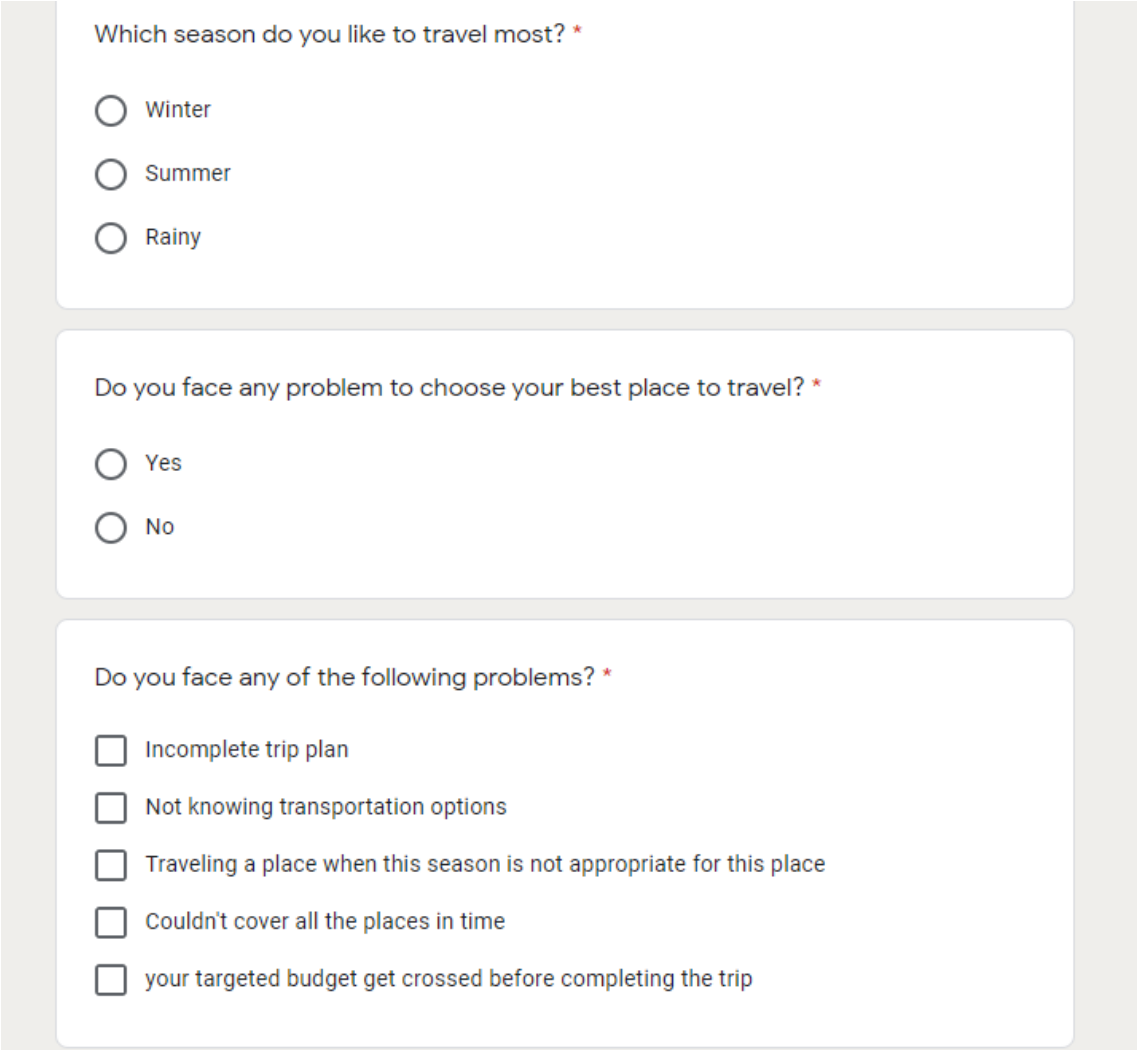


Figure 8.3: Part 3

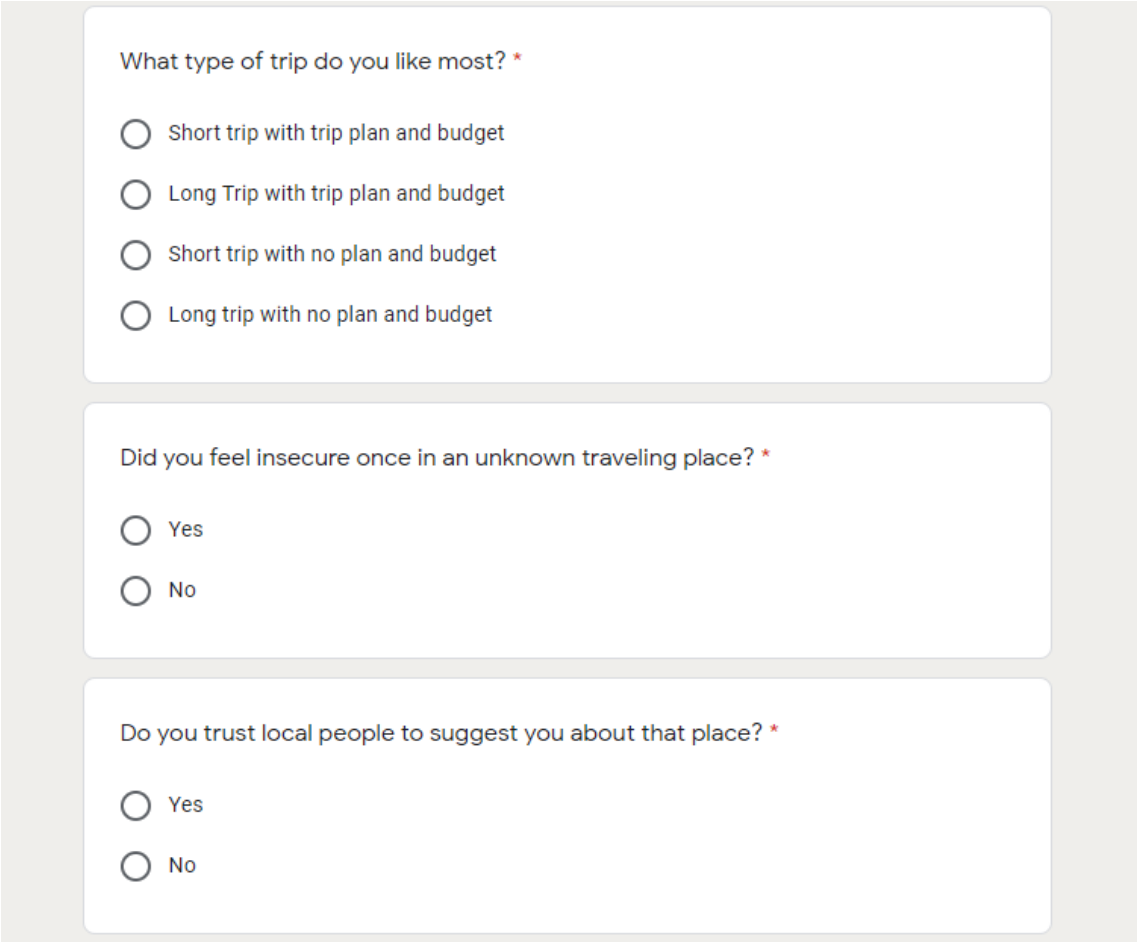


Figure 8.4: Part 4

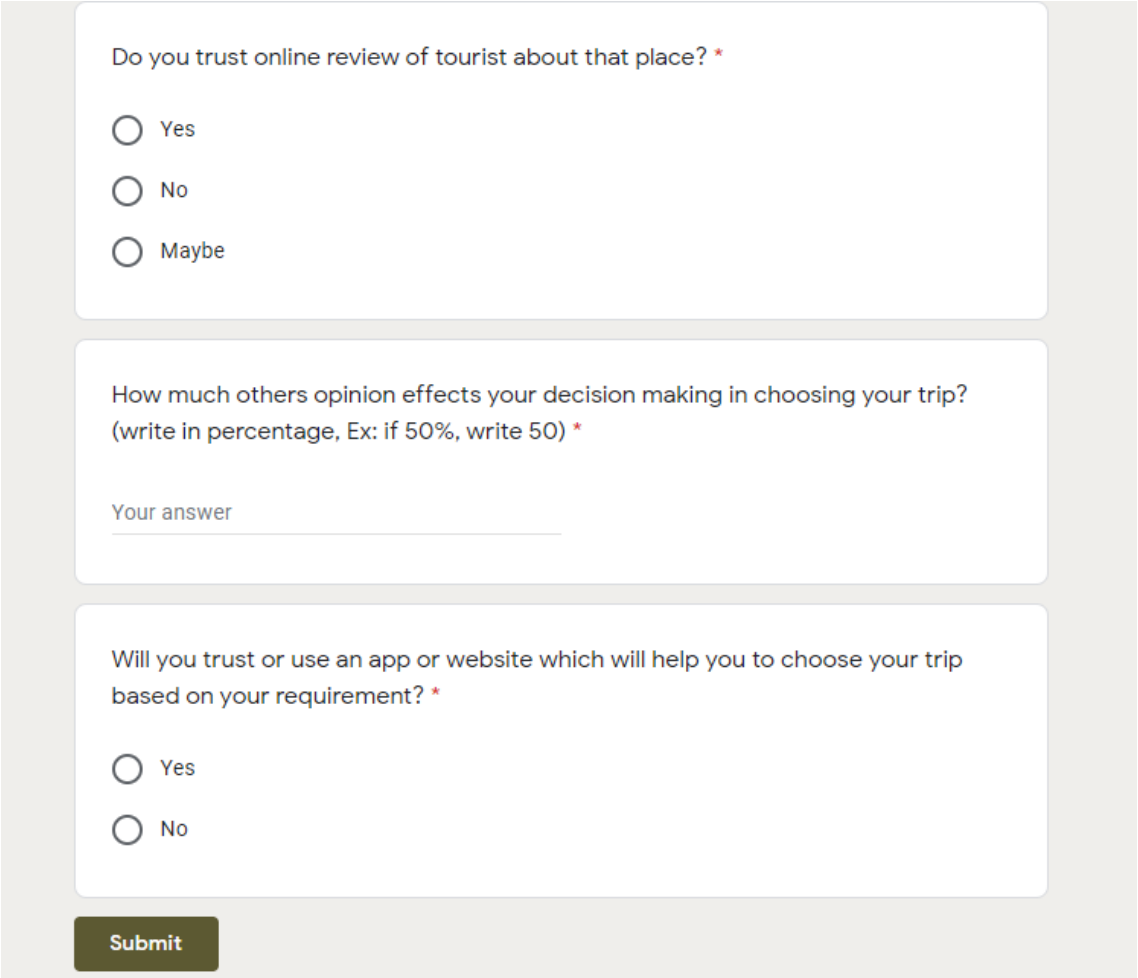


Figure 8.5: Part 5