

Nature Reserves Tourism System in Egypt Using Gis Case study

Ahmed Waleed , Eiad Fathy, Ahmed Mohamed , Mohaned Ahmed

Supervised By : Prof. Dr. Abdelnasser riad, Eng.MennatAllahHassan

March 7, 2020

1 Introduction

1.1 Purpose

The purpose of the software design document is to provide a full description of the architecture and system design of tourism management system and rating system. This document will explain the components of the system in the block diagram, illustrate how certain tasks done between system and the user with sequence diagram[1], the development and implementation of the project represented in class diagram also the explain the logical structure of the database represented in ER diagram.

1.2 Scope

Tourist will benefit from this application to get fully certified data about national park for free as it is done for an academic purpose also, the application will be available in many interface that will be available for almost all devices to cover the maximum number of users. Application will provide users with feedbacks of each park to make sure the tourist get a real feedback for is target national park to choose what he need the most from his trip.

1.3 Overview

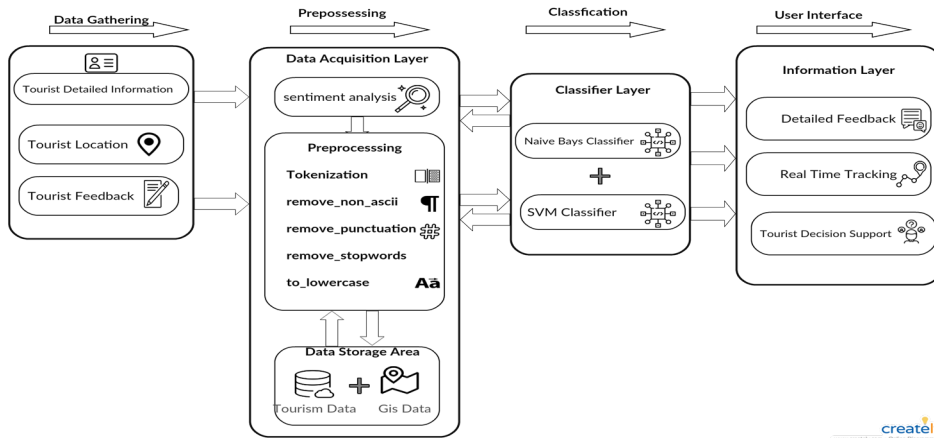
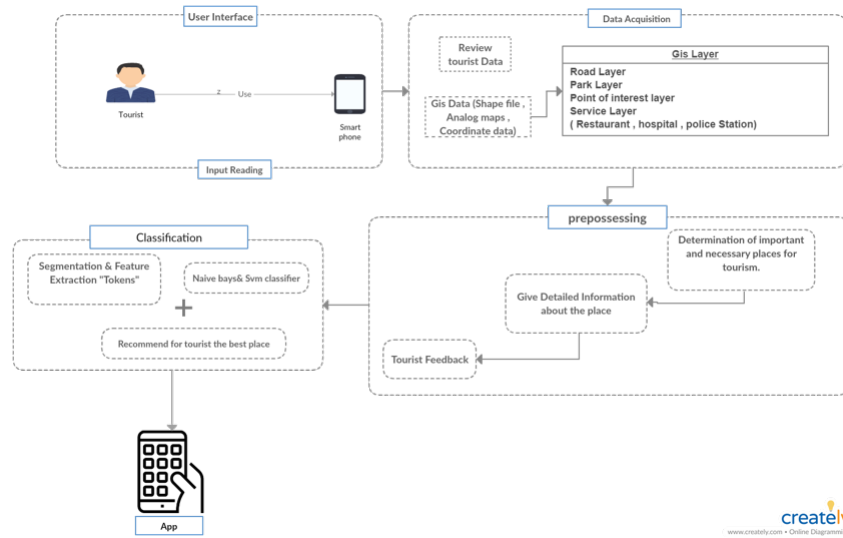
Tourists cannot access all Egyptian protected areas they can access only few of it. So, this system will help tourists to know all about protected areas in Egypt from its road and services around each of them. Rating system of protected areas aims to give tourists a trustworthy feedback of protected areas on many classes; which also will help and assist protected areas managers to improve it if possible and also it will assist tourists which protected area matches tourists. This document explains the use of computer science for tourism.

1.4 Definitions and Acronyms

Term	Definition
Software Design Document (SDD)	Used as the primary medium for communicating software design information.
NLTK	Natural Language Toolkit is a platform used for building Python programs that work with human language data for applying in statistical natural language processing.
NLP	Natural language processing is a subfield of computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human languages.

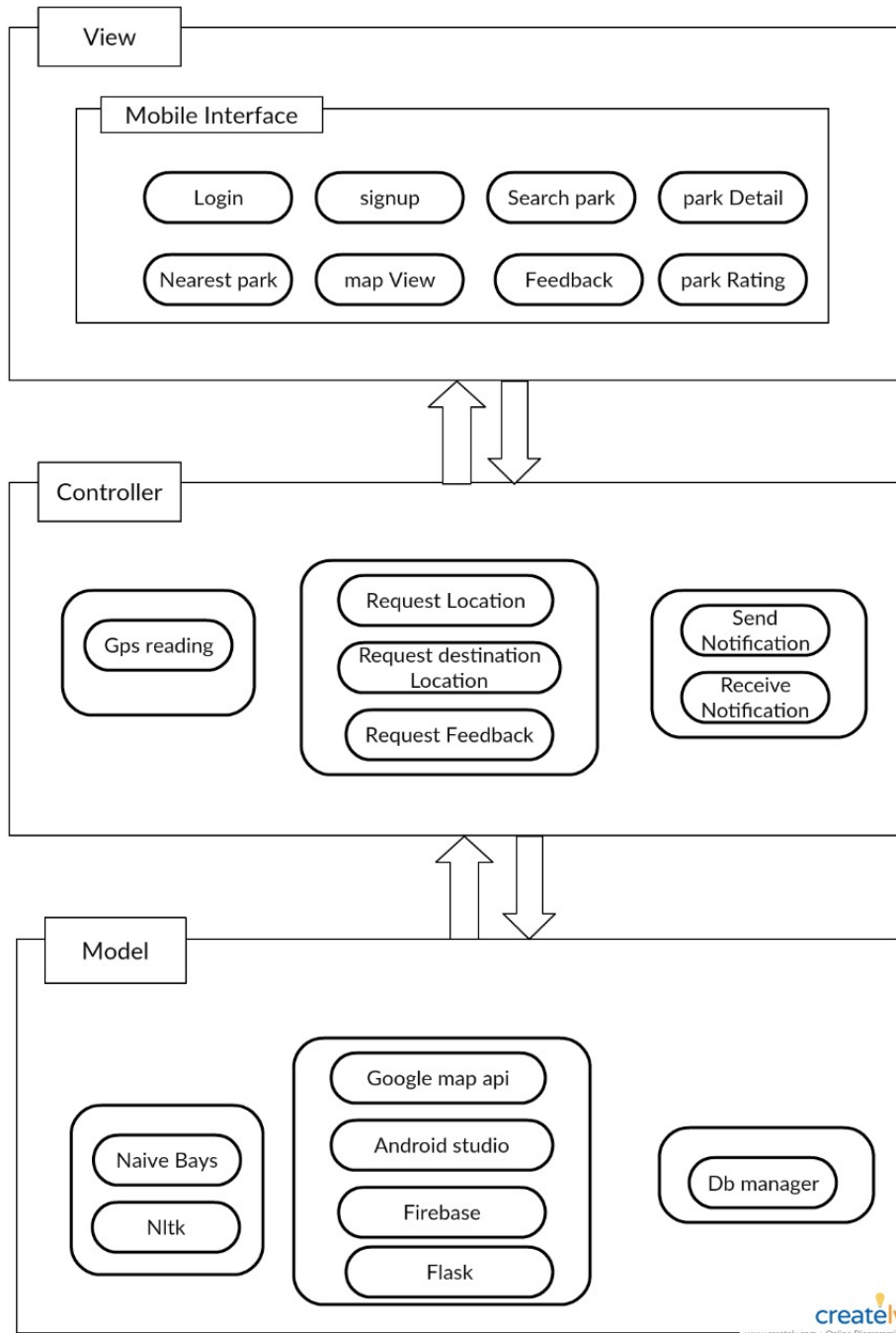
2 System Overview

Tourist will use only a smart phone to use application. In order to get a high accuracy of rating system of protected areas , The system has to make some tasks from classification and preprocessing, That the use of (svm, naïve bays) classifiers in rating system after (Tokenize, remove non ascii, remove punctuation, remove stop words, to lowercase) will increase the accuracy and use Geographic information system on developing maps that offers the ability to manipulate and maximize data, be it spatial. Also, there will be a business monitor to view detailed rating of any tourist, tourist can view park's location and rate park and give feedback.



3 System Architecture

3.1 Architectural Design



3.1.1 Model

The interpreter and expert models deal with simultaneous interpreters that will interact with the system either by using the main interpreting functionality or add and extract data from database.

3.1.2 View

It is responsible for presentation of data and information throughout the user interface (UI).

3.1.3 Controller

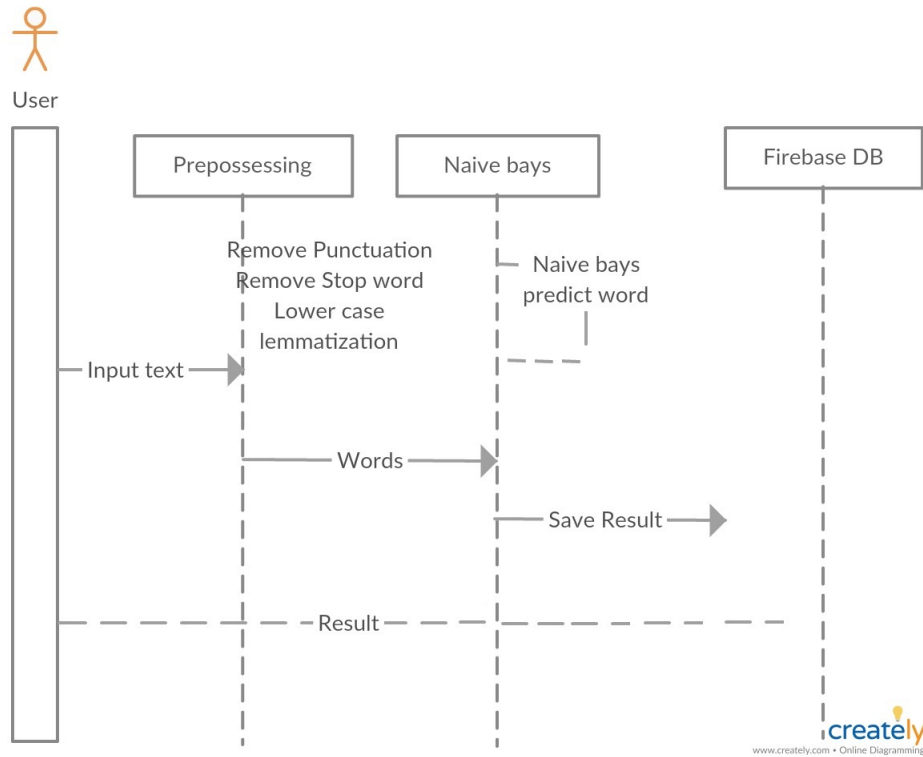
It is responsible for binding the view and model.the interactions and requests made within the view are taken and sent to the database to fetch data with the use of models then it forward data to the view again to be shown.

3.2 Decomposition Description

3.2.1 System Sequence Diagram

Each step that occurs across all subsystems to achieve text analysis is described in the figures below.

3.2.2 Classification Sequence



Rationale: The user can input text (Review) to the system in the form of plain text and receive a Rating from all classes the user review.

Input: Text (Review).

Output: Rating.

Priority: High. This is an important functionality in system.

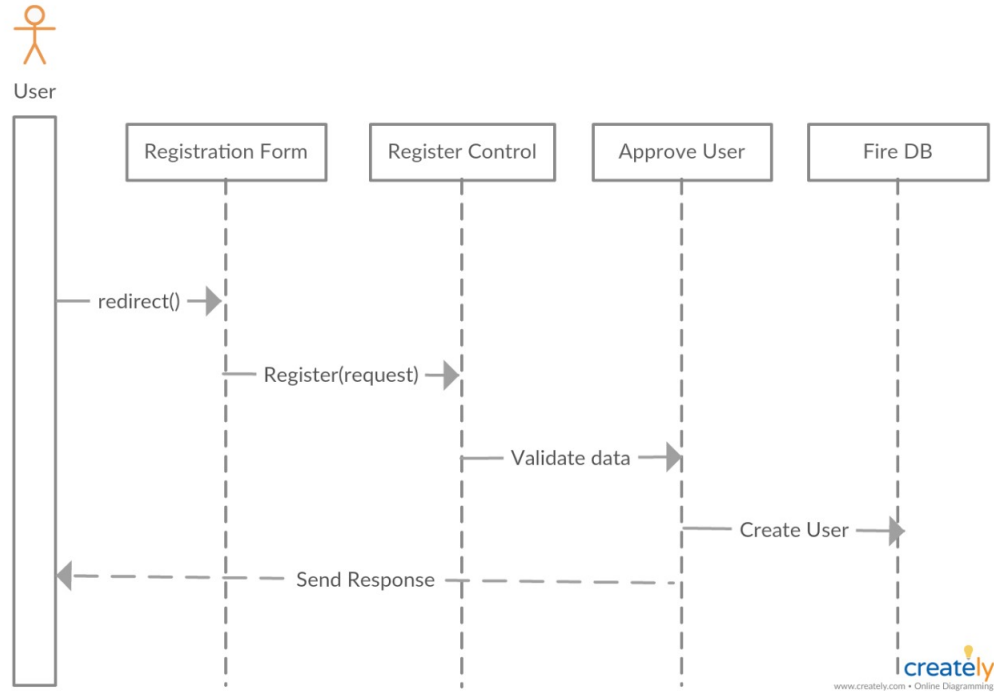
3.2.3 User Register Sequence

Rationale: The user should register for an account before being able to use the application.

Input: name, e-mail, password.

Output: Validation response.

Priority: Medium. Each user must have an account and become assigned a role before being able to use the application.



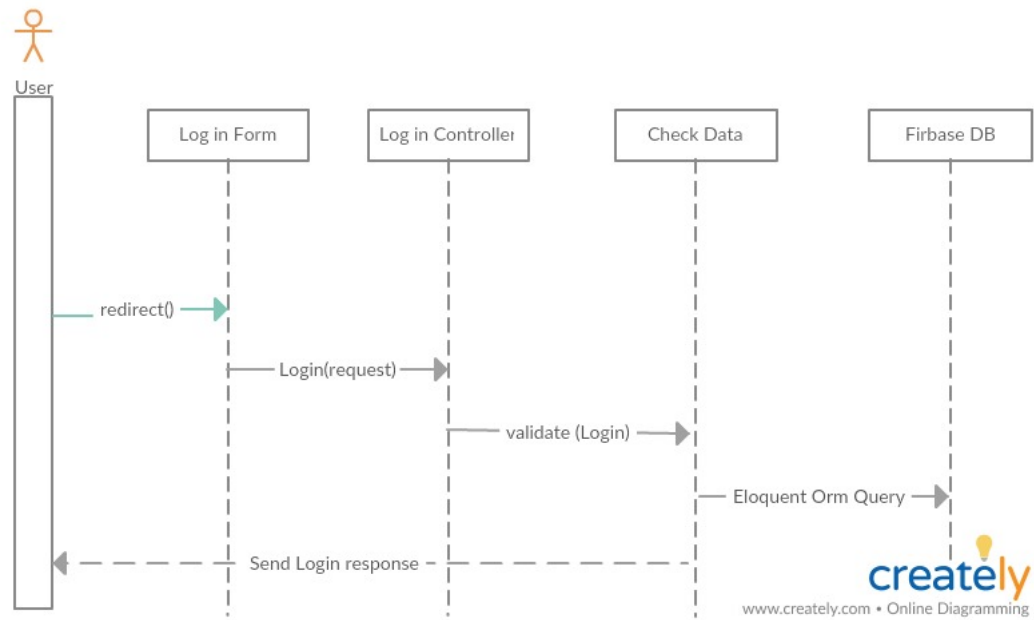
3.2.4 User login Sequence

The User should login to be able to enter-act with the system.

Input: name, e-mail, password.

Output: Validation response.

Priority: Medium. Each user must have an account and become assigned a role before being able to use the application.



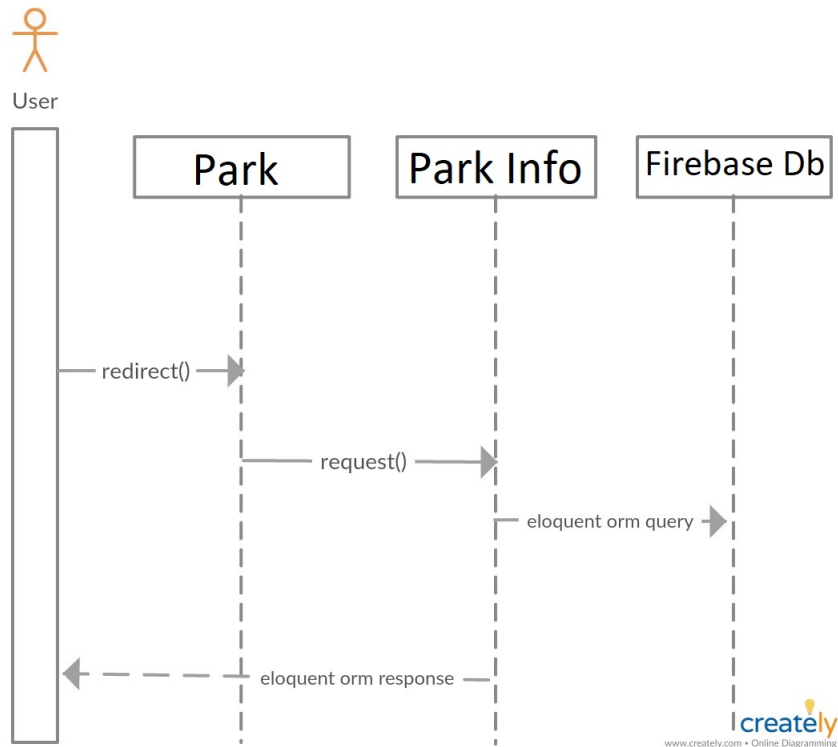
3.2.5 check parks Sequence

The system should be able to search and identify the national park that the tourists (User) want to go to it.

Input: park name, address.

Output: search response.

Priority: high.



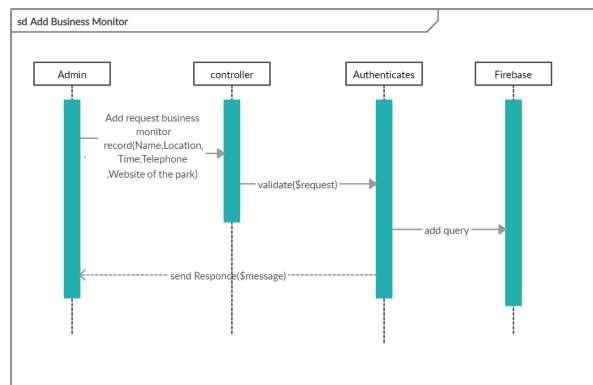
3.2.6 Add Business Monitor Sequence Diagram

Rationale: Admin can add Business Monitor record.

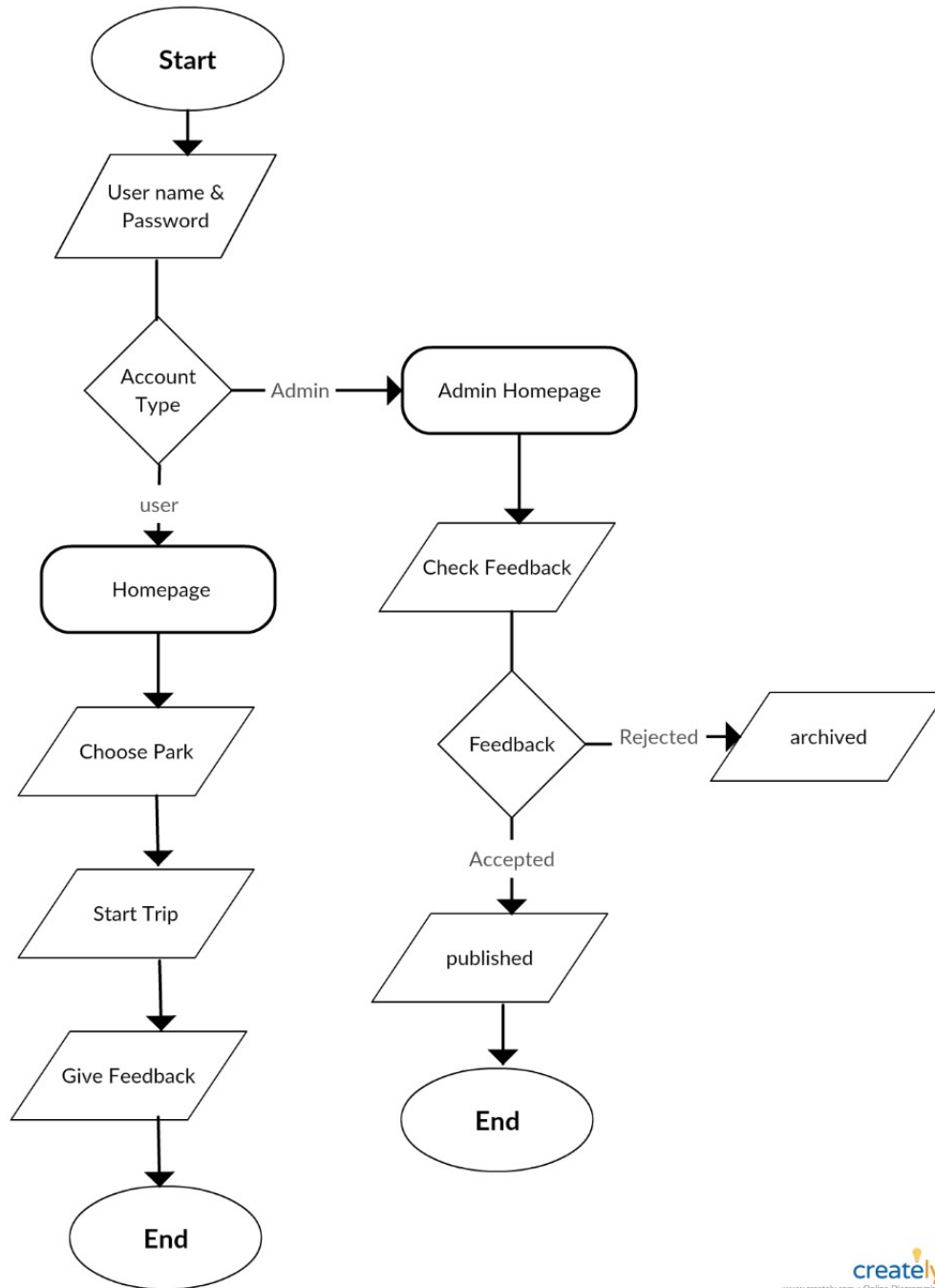
Input: Text (Name ,Location ,Time ,Telephone ,Website of the park).

Output: Response message if some thing went wrong or not then add new record into firebase.

Priority: High. This is an important functionality in system.



3.3 Process Diagram



3.4 Design Rationale

As mentioned previously, we have used MVC as our architecture as it helped us separate the functionality and data of our system from the presentation. In MVC design pattern, where the model (Cloud) handles the back-end processing like feature extraction and classification, the controller takes the information from the model and deliver it to the view for the user interface, the view takes the event from the user using the event listeners and deliver it to the controller, then the controller gives it to the model for further processing. So, we can easily make modifications, re-use and optimize functionality part as it is our core. Also the software we are developing efficiency and accuracy is a very important aspect of it so it will be very sensitive with data so this should be developed in a very accurate and efficient way.

3.5 Possible Algorithms

There were so many alternative algorithms that we could have used like decision trees and SVM that are generally very good for text classification. Naive Bayes: It works on conditional probability. It calculates the probability of the input relative to some decisions that was previously taken. It is well suited when the input has large number of dimensions. Decision Trees: It is a decision support tool that uses a tree-like model of 11 decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. SVM: Linear SVM is given a set of train data which belong to a certain class to find an optimal separating line. It tries to maximize the distance between each class to avoid misclassification. Then a test data are given to be classified to one of the classes formed before.

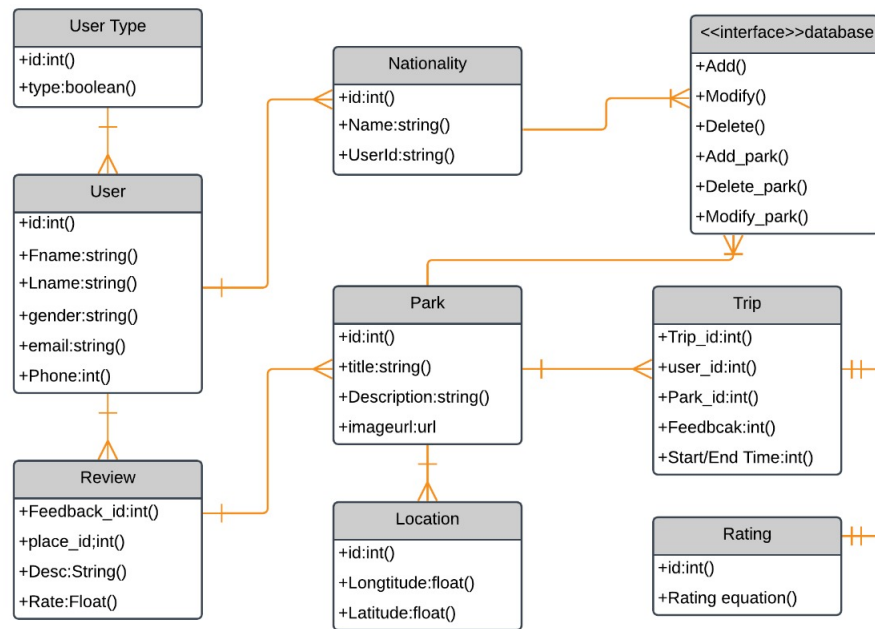
3.6 Frameworks

There are multiple frameworks that can be used like Django and Laravel and flask frameworks for PHP to make the transition from the python to the web app easy to maintain and scalabale. Flask: Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions. PHP Laravel: Laravel is a free, open-source PHP web framework, intended for the development of web applications following the (model -view -controller) architectural pattern. It was mainly used for the creation and main-taing of the web applicataion. Django RESTful API: Django REST framework is a powerful and flexible toolkit for building Web APIs. Django is a high-level Python Web framework encouraging rapid development and pragmatic, clean design. A web application framework is a toolkit of components all web applications need. Meaning that we with the help of flask we can integrate out python code which contains the machine learning part and text classification into the

web app with ease.

4 Data Design

4.1 Data Description



4.2 Data Dictionary

user	Attribute: id, name , password, country, email, phone number. This table hold all data of each user and create an unique id for each user.
reviews	Attribute: feedback id, place id, description, rate. This table holds all feedback given by an user for each park and generate a rating for each park
parks	Attribute: id, title, name, image url , rate, title This table holds all information for each national park available in the system and holds also the rating of each park
Rating equation	Attribute:id This table holds all equation and the equation to calculate the rating for each park according to the given feedback.
User type	Attributes id,type This table holds each user type either admin or normal user.
Database “interface”	This table holds all methods will be used by an admin user.
trips	Attributes: trip id,user id,park id, feedback This table holds each trip by every user .
Trip road	Attributes: id This table holds the road of each trip referred by an id .
road	Attributes: id This table hold each road on the map.
Location	Attributes: latitude,longitude,park id This table holds each park location.

5 Component design

-Naive bayes: Naive Bayes classifier is a straightforward and powerful algorithm for the classification task. Even if we are working on a data set with millions of records with some attributes, it is suggested to try Naive Bayes approach. Naive Bayes classifier gives great results when we use it for textual data analysis. Such as Natural Language Processing, as needed in our project.

-SVM: A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After

giving an SVM model sets of labeled training data for either of two categories, they're able to categorize new examples.

5.1 Pre-Processing

5.1.1 Remove Punctuation

cleans up the string and remove all the punctuation marks. removes all punctuation's from a string. punctuations: `"'!()-[];:'"i./?@"`

5.1.2 Remove Stop Words

Stop Words: A stop word is a commonly used word (such as “the”, “a”, “an”, “in”) that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query. We would not want these words taking up space in our database, or taking up valuable processing time. For this, we can remove them easily, by storing a list of words that you consider to be stop words. NLTK(Natural Language Toolkit) in python has a list of stopwords stored in 16 different languages.

5.1.3 Lemmatization

Lemmatization is the process of grouping together the different inflected forms of a word so they can be analysed as a single item. Lemmatization is similar to stemming but it brings context to the words. So it links words with similar meaning to one word. Text preprocessing includes both Stemming as well as Lemmatization. Many times people find these two terms confusing. Some treat these two as same. Actually, lemmatization is preferred over Stemming because lemmatization does morphological analysis of the words.

5.1.4 Lower case

changes all capital letters to small letters inorder to be transfered to the text classification method with out any errors.

5.2 Text Classification

After that we use the words to do text classification on the input text to extract the Rate of it. This was done by using the Naive Bayes Classifier. To complete the process we trained the Naive Bayes Classifier, it was used as it is the best Classification Method available working with text. The idea mainly is to create a model that can predict the the Rating of tourists using only words of the input text and to make the task of the Rating terms easier for users and tourists.

6 Humnan Interface Design

6.1 Overview of User Interface

Since there is a web interface for the system is for the business monitor end user, and an android based mobile application for the tourist's end user; therefore, we have two different user interfaces. Business Monitor Web Interface: The business monitor will be given his/her account to be able to log in. Once logged in, (s)he is redirected to the homepage which contains the users accounts and reviews. All web pages will be provided a navigation menu, too be able navigate through the website as desired. tourist's Android Application Interface: The tourists will be able to log in to his/her account and will be shown activity screen which will allow him/her to choose the park desired and then a route will be drawn from the starting destination to the end destination, there is also a menu in app to make the user more easily to navigate in the various activitys of the application, there are search for parks, see previous reviews of that park, see there comments, be able to rate and review for the whole aspects of the trip form the transportation to the food , weather and so on.

6.2 Screen Images

6.2.1 Android Application

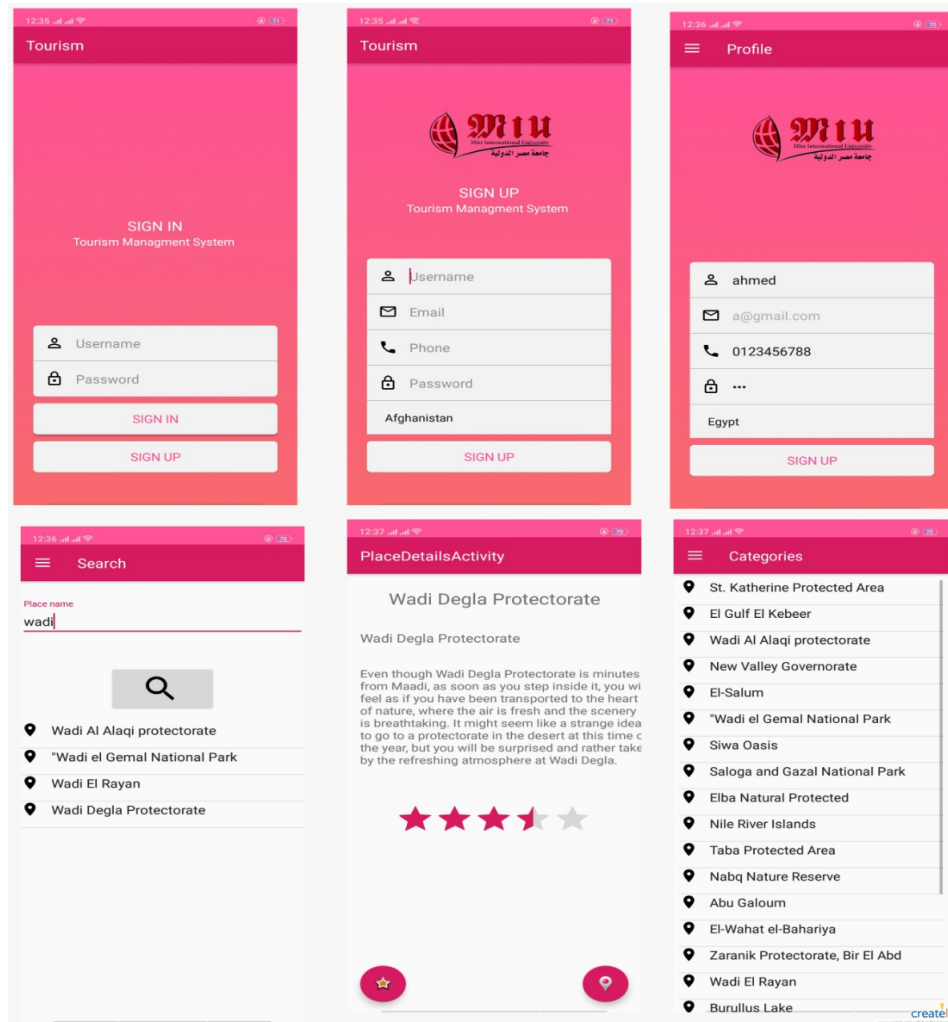


Figure 1 : Login

Figure 2 : Register

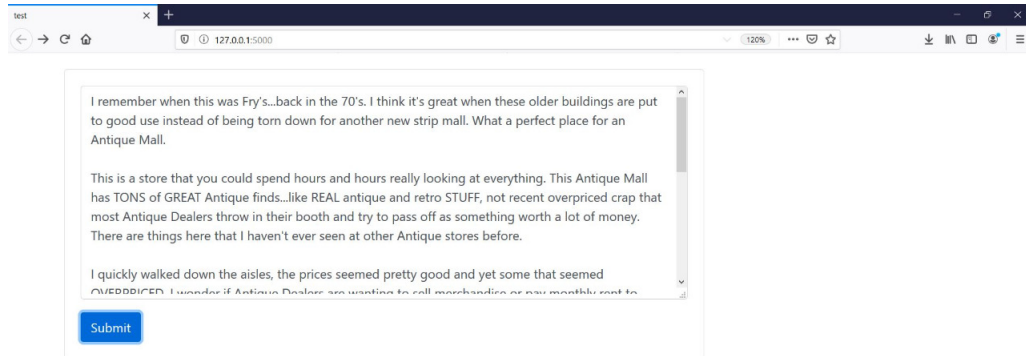
Figure 3 : Edit Profile

Figure 4 : Search for Park

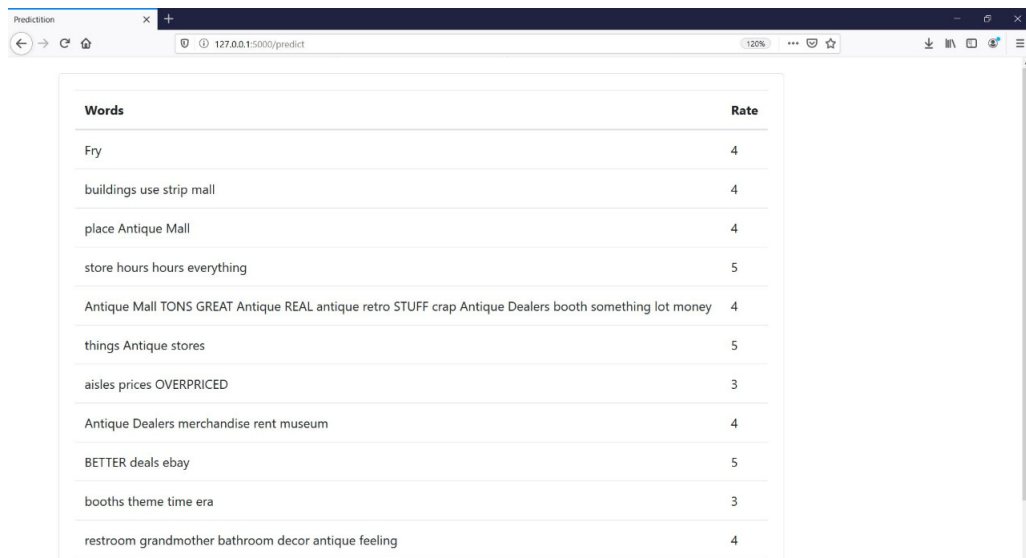
Figure 5 : Rate and Re Park

Figure 6 : Parks list

6.2.3 Web Application



in this web interface the text-box will be used as trial for the machine learning classification.



the review will be tokenized and split into words which will contain the rate of the user. it will be shown in form of table.

6.3 Screen Objects and Actions

- Search Button, to search for specific protected area.



Figure 1: search for protected area.

- This button show a full description of a protected area and its rating.

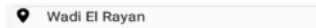


Figure 2: protected area button for each one.

- This button allows the tourist to view protected area location.



Figure 3: location button.

- This button allows the tourist to review a protected area.



Figure 4: review button.

- This button shows all protected area.

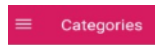


Figure 5: categories button.

7 Requirements Matrix

Requirements Traceability Matrix				
ID	Status	Functional Requirement	Technical specification	Implemented In
1	Completed	TokenizeText	removing all the spaces and the breaks in each review.	pre_processing
2	Completed	Classify Data	Perform Naïve bays and SVM classification to enhance rating system performance	
3	Completed	LowerCase To All	It take the array of words output and Convert into Lowercase	pre_processing
4	In Progress	Request Trip	The Tourist requests a Trip sending his/her current location which is sent to Admin	
5	Completed	lemmatization	It take the array of words output and get the single forms of derived words such as:("students ":"student")	pre_processing
6	In Progress	View Tourist's rating	The Admin views the rating of the Tourists	
7	Completed	RemoveStopWords	It takes the array of words output and removes any stopword such as :("as","the","a")	pre_processing
8	In Progress	Adding a Business Monitor	The admin adds a business monitor record to the database	
9	Completed	Remove Punctuation	It takes the array of words output and remove Punctuation	pre_processing
10	In Progress	Editing a Business Monitor's Account	The Admin edits a business monitor's account to the firebase	

8 APPENDICES

- ArcGIS: a geographic information system (GIS) for working with maps and geographic information. It is used for creating and using maps, compiling geographic data, analyzing mapped information, sharing and discovering geographic information, using maps and geographic information in a range of applications, and managing geographic information in a database. The system provides an infrastructure for making maps and geographic information available throughout an organization, across a community, and openly on the Web.
- GIS: A geographic information system (GIS) is a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data. The key word to this technology is Geography – this means that some portion of the data is spatial
- Google maps: Google Maps is a Web-based service that provides detailed information about geographical regions and sites around the world. In addition to conventional road maps, Google Maps offers aerial and satellite views of many places. In some cities, Google Maps offers street views comprising photographs taken from vehicles.
- Semantic analysis: is the process of relating syntactic structures, from the levels of phrases, clauses, sentences and paragraphs to the level of the writing as a whole, to their language-independent meanings. It also involves removing features specific to particular linguistic and cultural contexts, to the extent that such a project is possible. The elements of idiom and figurative speech, being cultural, are often also converted into relatively invariant meanings in semantic analysis. Semantics, although related to pragmatics, is distinct in that the former deals with word or sentence choice in any given context, while pragmatics considers the unique or particular meaning derived from context or tone. To reiterate in different terms, semantics is about universally coded meaning, and pragmatics, the meaning encoded in words that is then interpreted by an audience.
- SVM: is a supervised machine learning model that uses classification algorithms for two-group classification problems. So you're working on a text classification problem. You're refining your training data, and maybe you've even tried stuff out using Naive Bayes.

9 References

[1]Kayoko YAMAMOTO, Tsukasa IKEDA. “Social Recommendation GIS for Urban Tourist Spots”. In 2016 IEEE 18th International Conference on High Performance Computing and Communications; IEEE 14th International Conference on Smart City; IEEE 2nd International Conference on Data Science and Systems.