Kathmandu Bernhardt College

(Affiliated to Tribhuvan University)

A Project Proposal

On

"Memories Finder Using Spherical Law of Cosine and Naïve Bayes"

Under the supervision of

Dabbal Singh Mahara

Submitted by

Kapil Bhattarai (5070/071)

Kishan Khadka (5071/071)

Sudip Karmacharya (5086/071)

Submitted to

Department of Computer Science and Information Technology

Kathmandu Bernhardt College

Bafal, Kathmandu, Nepal

April, 2018

1. Introduction

Memories Finder is a Location-Based Social Networking Android application intended for those people who like to leave their spent memories in the form of photographs or leave reviews to the world, whenever, they visit places. System shows nearby Point of Interests (POIs) and Business Point of Interests (BPOIs) of previously visited place by their friends within 10km radius from user's current location. The System recommends top 10 nearby places based on sentiment analysis of reviews and rating from people. Additionally, Users can get information of nearby public events like blood donation, awareness program, sporting program, exhibition, street program and many other important activities.

2. Problem Statement

Many people are having problem to search proper places due to communication overhead or less facility of location guide. It is impractical to search each and every places at every location. Further, one might go to some places, upload status on social network and after some days, those memories fade away eventually because of flood of status on social network these days. It's hard to find which of ours friends visited that place before and see their memories at right time. Most of the time, people can't attend public events happening around them due to lack of prior information about it.

3. Objectives

The objectives of this application can be summarized as:

a. To remind memories of friends and families with user's current location
b. To recommend nearby places to users
c. To provide information about nearby social events
d. To provide location information to traveler

4. Scope and Limitations

4.1. Scope

Android is a leading mobile operating system dominating rest of the mobile operating system. Memories Finder is an android application which allows users to see memories of their friends within 10km radius from their current location. System recommendations top 10 places which helps travelers to choose the perfect nearby location to visit. With this system, user can share their memories with the world using android smartphone and stay

connected with friends and family members and their memories as they travel to different places. Users can use this application to get information about nearby upcoming social events which helps to plan a schedule for social contribution out of their busy work life. Therefore, this system has very good scope.

Table 4.1 Worldwide Smartphone Sales to End Users by Operating System in 1Q17 (Thousands of Units) [1]

Operating System	1Q17 Units	1Q17 Market Share (%)	1Q16 Units	1Q16 Market Share (%)
Android	327,163.6	86.1	292,746.9	84.1
IOs	51,992.5	13.7	51,629.5	14.8
Other OS	821.2	0.2	3,847.8	1.1
Total	**379,977.3**	**100.0**	**348,224.2**	**100.0**

4.2. Limitation

Different limitations of this project are:

a. Users are supposed to have Android phone with GPS support
b. Application needs user's location permission to share their memory

5. Study of Exiting System

Foursquare is a Location Based Social Networking Application with over 50 million users worldwide. It is a local search-and-discovery service mobile app which provides search results for its users. The app provides personalized recommendations of places to go to near a user's current location based on users' "previous browsing history, purchases or check-in history". DBSCAN (Density-Based Spatial Clustering of Applications with Noise) and Multi-Output random Forest algorithm is used to improve map pins [2]. Sentiment Analysis of users' reviews is taken into consideration as major parameter for recommendation.

Nearest tourism site searching using Haversine method was studied [3]. 13 tourism sites of Samarinda city, Indonesia were used. The method used to search the nearest tourism site based on the user's location is Haversine method

Classifying sentiment in Arabic social network using Naïve Bayes was studied. Two different approaches were used to classify Arabic Facebook posts. The first one depends on syntactic features, using common patterns used in different Arabic dialects to express opinions. The highest coverage and accuracy achieved in this approach were 49.5% and

83.4 % respectively. The second approach is a probabilistic model, Naïve Bayes classifier. The highest coverage and accuracy achieved in this approach were 60.5% and 91.2% respectively [4].In order to determine the users nearby locations, Great Circle Distance between the user's current location and places is needed. There are 3 known major algorithms to compute Great Circle Distance:

i. Haversine:

Haversine formula is used for Navigational purpose, it gives great circle distance between two points on a sphere based on their latitude and longitude [5].

ii. Vincenty

Vincenty's formulae are two related iterative methods used in geodesy to calculate the distance between two points on the surface of a spheroid, developed by Thaddeus Vincenty. They are based on the assumption that the figure of the Earth is an oblate spheroid, and hence are more accurate than methods such as great-circle distance which assume a spherical Earth [6].

iii. Spherical law of Cosine

For a spherical triangle with surface angles A, B, C opposite sides corresponding to central angles a, b, c, the spherical law of cosines states [7] ,
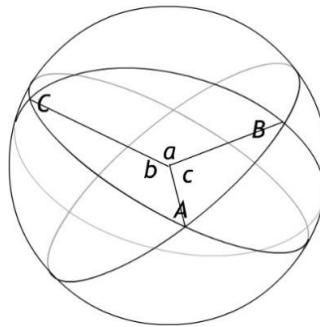
$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(C)$$

Figure 5.1: Spherical law of Cosine
[Source:" http://www.math.stonybrook.edu/~tony/archive/hon101s08/spher-trig.html"]

Among three, Spherical Law of Cosine will be used because it is faster than Haversine and Vincenty Algorithm [8]. Haversine algorithm produces much accurate output compared to Spherical law of Cosine below 1 meter radius [9]. However, system is considering 10 km radius from user's location, Spherical Law of Cosine and Haversine both produce same results. Considering efficiency, Spherical Law of Cosine is suitable for the system.

For Sentiment Analysis of reviews, Naïve Bayes classifier will be used because it is faster than Support Vector Machine, Decision Tree and Bayesian Network and produces high accuracy [10]. It is a simple probabilistic model based on the Bayes rule along with a strong independence assumption. That is, given a class (positive or negative), the words are conditionally independent of each other. This assumption does not affect the accuracy in text classification but makes it really fast classification algorithms [11].

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

6. Requirement Collection

6.1. Functional Requirement

i.   User Registration:

Registration is open to all people. Users can register to Memories Tracker using Google or Facebook account.

ii.   User Function:

There are two types of users: Registered and Non Registered User. Registered users have full access to features of Memories Tracker. Non Registered Users cannot access features such as upload memories, rate place, like and comment to other memories. They can only view locations and public memories of people.

iii.   User Profile:

Application allows the registered users to access their own personal profile where user can include their different personal information.

6.2. Non Functional Requirement

i.   User Friendly:

 Users with basic knowledge social networking application can easily use. Memories Finder uses a simplified design and navigation, as well as simple language to improve the user friendliness of the application

ii.   Easy Access:

Memories Finder is mobile application. Thus, it can be accessed anytime from anywhere with application installed mobile and Internet Connection.

iii.   Speed of Application:

The speed of application depends on basically two factors: GPS Support and Internet Speed.

7. Feasibility Study

7.1.   Technical Feasibility

This android mobile application is technically feasible. The software and hardware requirement for the development of this application are not many and already available as free as open source. The work for the project is done with the current equipment and existing software technology. This application is supported by all android application having version 5.0.

7.2.   Operational Feasibility

The application will be implemented in android platform as there are more than 2.32 billion active android smartphones so the users can use this application in their existing android smartphone. Also the application will be developed using android SDK (27.0.0) so every android smartphones with Android Oreo or lower is supposed to run the system smoothly. Hence, the System will be operationally feasible.

7.3.   Economic Feasibility

This application will be economically feasible because application is made using open source software.

7.4.   Schedule Feasibility

Estimated time spent in the development of this application is 4 months and scheduled as:

ID	Task Name	3/25	Apr 2018				May 2018				Jun 2018				Jul 2018		
			4/1	4/8	4/15	4/22	4/29	5/6	5/13	5/20	5/27	6/3	6/10	6/17	6/24	7/1	7/8	7/15
1	Planning	████																
2	Requirement Analysis		████															
3	Feasibility Analysis			████														
4	System Design				████													
5	Implementation						████											
6	Documentation	████████████████████████████████																

Figure 5.4.4: Project Schedule

8. Data Collection

8.1. Data Collection Methodology

a. Databases and Records

Test Data for location includes latitude, longitude and name of the POIs and BPOIs from Geo Database called Geonames. Snapshot of the Test Data is shown below:

Table 8.1: Test Data for Latitude and Longitude

Name	Latitude	Longitude
Tribhuvan University	27.680677843664434	85.28347492218016
Kathmandu Darbar Square	27.7042755896117	85.30733585357667
Swoyambhunath Stupa	27.714761839072718	85.29038429260254
DakshinKali Mata Temple	27.67731683995404	85.28049230575563
Damascus, Syria	33.513808727209145	36.2765496969223
New Dehli, India	28.614013938417106	77.2093217074871
Dallas City Hall	32.77617748193533	-96.79733991622925

Test Data for Sentiment Analysis is collected from Senti word [12]

Table 8.2: Test Data for Sentiment Analysis

Positive Words	Negative Words	Neutral Words
Awesome	Bad	Idea
Good	Disliked	Hmm
Peaceful	Worst	Judgment
Perfect	Bore	Impact
Brilliant	Fake	Suppose

9. Tools

9.1. Frontend Tools

9.1.1. Android SDK

Android software development kit (27.0.0) contains the necessary tools to create, compile and package Android applications. The android SDK contains required libraries, debugger, an emulator, android debug bridge for connecting virtual or real device and relevant documentation. Android SDK 27.0.0 will be used for the development of application.

9.1.2. Android studio (IDE)

Android Studio (3.1) is the official integrated development environment for Google's Android operating system. Android Studio will be used as an IDE.

9.1.3. Android Virtual Device (AVD)

An android Virtual Device (AVD) is an emulator configuration that allows developers to test the application by simulating the real device capabilities. Google nexus 5.5 will be used as android virtual device.

9.1.4. Java

Java is a very popular object oriented programming language developed by Sun microsystems. In the development of the android application, java will be used as the main programming language.

9.1.5. XML

XML stands for Extensible Markup Language which is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable.XML is used to display and manage the GUI components. All the layout of the project will be defined by the XML. Besides these, defining the animation, color, style, string, dimension, menu etc. will be done using XML.

9.2. Backend Tools

9.2.1. MySQL

MySQL (5.7.21) is a freely available open source Relational Database Management System (RDBMS) that uses Structured Query Language (SQL). MySQL will be used as our database because it provides complex querying capabilities over firebase.

9.2.2. Apache Server

Apache Web Server (2.4.33) is open source Web server creation, deployment and management software. Apache server will be used to host PHP scripts.

9.2.3. GitHub

GitHub is a version controlling system. GitHub will be used for code sharing among group members.

9.2.4. Composer

Composer (1.6.3) is an application-level package manager for PHP that provides a standard format for managing dependencies and required libraries.

9.2.5. Slim

Slim (3.9.2) is a PHP micro framework that helps users quickly write simple yet powerful web service applications and APIs. Slim will be used to build APIs.

9.2.6. Firebase

Firebase is a google based cloud service for making application. Firebase authentication, push notification and firebase analytics will be used.

10. Testing Techniques

Black box testing will be used to check the features of system. Functionality of the system will be checked by providing appropriate inputs and results will be compared with the expected results.

i. Sign In Test Case

Table 10.1: Sign In Test Case

Test Case ID	Test Scenario	Test Steps	Expected Result
TC-1	Check sign in activity with valid data	1) Open App and go to sign in activity. 2) Fill up the form 3) Click sign in button.	Redirect user to the main activity of the App
TC-2	Check sign in Activity with invalid data	1) Open App and go to sign up activity. 2) Fill up the form with invalid data. 3) Click sign in button.	User Should get Respective Error Messages

ii. Post Test Case

Table 10.2: Post Test Case

Test Case ID	Test Scenario	Test Steps	Expected Result
TC-3	Check new post done by users with valid data	1) Fill up the form with valid data 2) Click post button	User should get post successful message and redirected to home activity.
TC-4	Check new post done by users with invalid data	1) Fill up form with invalid data 2) Click post button	User should get respective Error Messages

iii.   Sentiment Analysis Test Case

Table 10.3. Sentiment Analysis Test case

Test Case ID	Test Scenario	Test Steps	Expected Result

TC-6	User gives review of a place	1) An Algorithm converts the review to points using Semantic Analysis.
		  i) If review is positive then positive rating should be inserted for a place in database.
		  ii) If review is negative then negative rating should be inserted for a place in database.
		  iii) If review is neutral then neutral rating should be inserted for a place in database.
			User should get a message that review was successful.

iv.   Logout Test Case

Table 10.4: Logout Test Case

Test Case ID	Test Scenario	Test Steps	Expected Result
TC-5	Logging out from the App	1) User press Logout button	User should be redirected to sign in page and session should be destroyed

11. High Level Design of Proposed System

When application will be started the system will check if the user is new user or not. If he/she will be a new user then they will be redirected to login screen from which they can login or continue the app without login. The user will be recommended places based on people's review. The user can view people's memories within 10 km radius from their current location. If the user is logged in then they can upload, like and comment the memories.
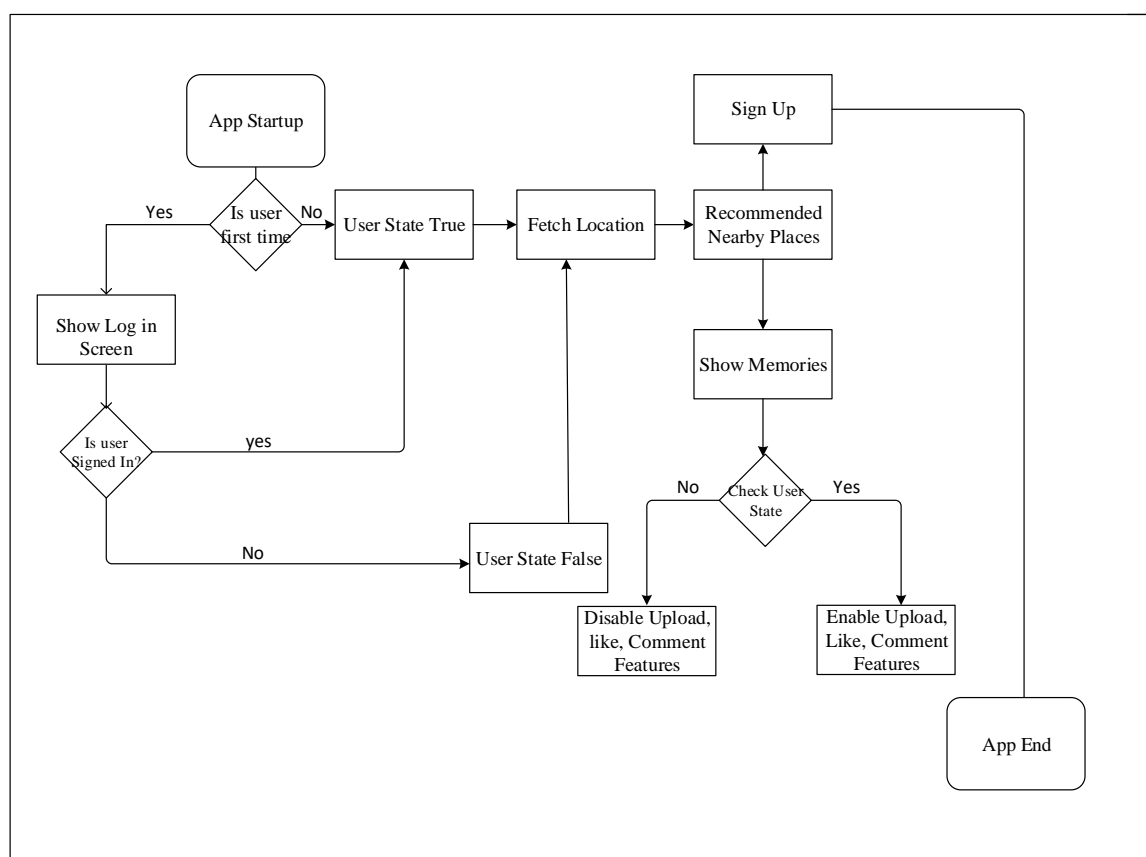

Figure 6.1: System Flow Chart

12. Expected Outcome

After the successful completion of this project, application will be able to recommend user's nearby point of interests and Business point of interests where they can see their friends memories and publicly uploaded memories. System will recommend nearby places based on Sentiment Analysis of user reviews and rating of places. User can upload their memories in terms of images and status by tagging geo location. Users will be able to use this application for navigation purpose while travelling and keep tracking of their friend's memories.

References

[1]   A. A. Forni and R. Meulen, "Gartner," 23 May 2017. [Online]. Available: https://www.gartner.com/newsroom/id/3725117

[2]   L. Xing, Z. Fan and R. Maddimsetty (2018, Mar. 19). You Are (Probably) Here: Better Map Pins with DBSCAN & Random Forests [online] Available:

 https://engineering.foursquare.com/you-are-probably-here-better-map-pins-with-dbscan-random-forests-9d51e8c1964d

[3]   Z. Arifin, M. R. Ibrahim and H. R. Hatta, "Nearest tourism site searching using Haversine method," *2016 3rd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), Semarang*, 2016, pp. 293-296.

[4]   M. M. Itani, R. N. Zantout, L. Hamandi and I. Elkabani, "Classifying sentiment in arabic social networks: Naïve search versus Naïve bayes*," 2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Beirut*, 2012, pp. 192-197.

[5]   D. S. AbdElminaam, H. M. Abdul Kader, M. M. Hadhoud and S. M. El-Sayed "Increase the Performance of Mobile Smartphones using Partition and Migration of Mobile Applications to Cloud Computing," *International Journal of Electronics and Information Engineering*, [On-line] Vol.1, PP.34-44, Sept. 2014.Available:http://ijeie.jalaxy.com.tw/contents/ijeie-v1-n1/ijeie-v1-n1.pdf#page=37 [Apr 4, 2018]

[6]   C. F. Karney, "Algorithms for geodesics", *Journal of Geodesy*. [On-line]. 87(1), pp.43-55 Available: https://link.springer.com/content/pdf/10.1007%2Fs00190-012-0578-z.pdf

[7]   G. V. Brummelen. "The Modern Approach: Oblique Triangles" *In Heavenly Mathematics: The Forgotten Art of Spherical Trigonometry.* (1st edition). Princeton University Press, 2012, pp. 94-110.

[8]   *Vincenty vs Haversine vs Spherical Law of Cosines Distance Calculations* [online].Available:https://jsperf.com/vincenty-vs-haversine-distance-calculations/5

[9]   D. S. AbdElminaam, H. M. Abdul Kader, M. M. Hadhoud and S. M. El-Sayed "Increase the Performance of Mobile Smartphones using Partition and Migration of Mobile Applications to Cloud Computing," *International Journal of Electronics and Information Engineering*, [On-line] Vol.1, PP.34-44, Sept. 2014.Available:http://ijeie.jalaxy.com.tw/contents/ijeie-v1-n1/ijeie-v1-n1.pdf#page=37 [Apr 4, 2018]

[10]   J. H. Jingjing, L. Charles and X. Ling. "Comparing Naive Bayes, Decision Trees, and SVM with AUC and Accuracy" .*Third IEEE International Conference on Data Mining, 2013*

[11]   V. Narayanan, I. Arora and A. Bhatia. "Fast and Accurate Sentiment Classification Using an Enhanced Naive Bayes Model". *Yin H. et al. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2013*, vol 8206, 2013

[12]   B. Liu, H. Minqing and J. Cheng. "Opinion Observer: Analyzing and Comparing Opinions on the Web."*International World Wide Web conference*, May 14, 2005.