Xenia – A Smart Tour Planning And Recommendation Using Crowdsourced Data

Ashish Pandey
Computer Engineering Department
Shivajirao S. Jondhale College of Engineering
Dombivli, Maharashtra, India

Raj Nikam
Computer Engineering Department
Shivajirao S. Jondhale College of Engineering
Dombivli, Maharashtra, India

Vikrant Nalawade
Computer Engineering Department
Shivajirao S. Jondhale College of Engineering
Dombivli, Maharashtra, India

Dr. Savita Sangam
Associate Professor/Head of Department of Information Technology Engineering
Shivajirao S. Jondhale College of Engineering
Dombivli, Maharashtra, India

***Abstract*—Tourism is one of the most important sectors to improve for any country that wants to develop its economy. Tourism nowadays is promoted by advertisements and served better using mobile and web applications. One such major service is Tourism apps which promotes tourism by providing users with services like travel booking, stay accommodations and information about the tour. But as world is diverse with languages and education this tourism app has difficulties reaching all users and serving them to the best. This study is going to address such problems and also provide a system that will solve these problem by an application that uses technologies like AI, NLP and recommender systems to solve such problems. The main goal of creating this app is to provide the user the right recommendation and make their tour planning sophisticated rather confusing. Also, this project is going to provide business to the locals in exchange for precious data of that particular place.**

I. INTRODUCTION

Mobile applications now play an important role in all areas of business and privacy. The travel and tourism industry is no exception because a large part of its success is based on these mobile applications. Paper maps, guidebooks, and other items are no longer in use, having been replaced by interactive and fantastic smartphone applications for the travel and tourism industry that offer a variety of features. This app provides expert guidance for the tourists in planning their tour ahead of the time. This app includes a smart recommender system which fetches the data from different tourists and their reviews and recommends the best location, restaurants and accommodation to the user. Also this app will help the local businesses to grow their business by simply providing their precious experience about the place to us. This crowdsourced data will be most accurate and this app will then be able to recommend the customers the best choices which totally depend on facts and experience.

II. LITERATURE SURVEY

P. Mate and H. Chavan [1], they proposed framework is an Android based Mumbai City Guide application which intended to handle area put together persistent inquiry with respect to the street organization. Someone when visit places, for example, Hotels, Colleges, Hospitals, and Schools don't have to employ an exceptional individual who give direction. On the off chance that all the data should be accessible on a cell phone with the client altered configuration, at that point it's useful to deal with their significant time successfully and effectively.

S .Bhattacharya [2], proposed a methodology that a local escort application called Mobile Campus on android based versatile stage for SRM University grounds. Close to handle correspondence (NFC) is a bunch of guidelines for PDAs and comparable gadgets to set up radio correspondence with one another by contacting them together or bringing them into closeness, typically close to a couple of centimeters. This local area expert application incorporates usefulness like finding current area of clients, showing college grounds map, course heading of college transport and gives little depiction and contact data of significant puts nearby.

Brown & Chalmers [3] composes that "sightseers intentionally make arrangements that are not exceptionally organized and explicit, so they can exploit evolving conditions." Although the arranging of conduct among voyagers shift this is profoundly pertinent for versatile travel applications. A voyager that makes definite arrangements has a lesser requirement for data "at this very moment" contrasted with explorers that do some arranging or practically no arranging in regard to what to do while they are at the objective.

U. Thakur [4], they proposed tools like expanded reality (AR) hold a tremendous potential in pulling in and holding guests. The ascent in shrewd cell phones just lifts this further as it gets conceivable to have data and visit age readily available.

Nowadays' cell phones are utilized to give various capacities notwithstanding ordinary voice correspondences. The capacities of advanced mobile phones empower the area based increased reality benefits a reality as referenced by J. Choi, B. Jang, G. J. Kim [5].

R. A. Abbaspour and F. Samadzadegan [6] built up a Context-Aware versatile local escort framework dependent on help situated architecture(SOA). PTG sends a solicitation dependent on the setting of client to a list. Administration merchants work together to track down the correct administrations and afterward, PTG and specialist organization haggle as to arrange of the solicitation and some other convention issues.

Tourism is the biggest business in the worldwide world economy utilizing around 200 million individuals and serving 700 million sightseers around the world. By 2020, the quantity of vacationer appearances all throughout the planet is required to increment by more than 200%. The travel industry is additionally answerable for creating an expected 11% of the worldwide (GDP) [7]. Despite the fact that, there is incredibly advanced travel data gave to the vacationers on the Internet and through the applications. Notwithstanding, there is no application for a vacationer to straightforwardly get the set of

experiences or some other data related with any landmark or spot by its image. To find out about any spot or landmark the vacationers visiting that spot need to utilize guides [8].

Mobile Tourism is a term that begins to show up over the most recent twenty years. It includes utilizing cell phone as electronic local area expert. The vacationer needs to look through data about a Point of Interest (POI) from his versatile. He can get the data from the web through an internet browser, yet this requires a consistent remote association with the web.

Overall the writing on explorer data needs and utilization of portable applications and gadgets in situ is dispersed and divided. Nonetheless, a few if the examination on data sources in the arranging stage is likewise valuable in situ. In the arranging stage a portion of the data search relates the decision of objective, how to arrive and convenience.

This study first attempt to enroll every one of the limits and difficulties experienced while using ideas of NLP to build up a local area expert framework. This model depict different best in class NLP and ML applications that offer such support, having their own arrangement of disadvantages, and give a concise prologue to proposed framework.

III. Proposed Approach

Tourism has become one of the major sector of revenue generation in the country, as promoting tourism has a direct impact on a country's economy. In the era of AI and WEB 2.0 tourism takes advantage of technology such as mobile applications and web applications, but not all users out there has the skill to operate this app to the fullest one such problem is lack of good customize recommendation system, and as the user database increases day by day people need an app robust enough to serve them. Also, need some application which will help the local businesses to grow their business by providing the useful data and experience to the customer. This data will eventually be recommended by recommended system to the users.

A. Phase 1

Xenia is a tour recommendation app that will help users to book tours and guide them through the tour. This app takes into account of the user's preference and suggests yours by using a recommender system. The user can then choose the tour based on the information that is provided like tour expense, ratings, review, the best time to visit, images and videos, and a short description. The design of this app have a user-friendly UI and run smoothly. Users can authenticate this app by creating a local account or using google sign in, this way it provides each user their personalized recommendations.

After choosing the tour the user can then navigate to the tour and choose "Start tour", then virtual tour guide will direct them by using live interactive maps, where this app provides information of each check point within the tour packages, and the user can then visit these checkpoints and complete the tour.

This app also provide a shopping tab that users can use to shop in the place where they visit, this tab will show the user's item based on the places they visit, E.g. : local products in Manali during their stay. This app uses flutter and dart programming for building UI and logics.

B. Phase 2

This app is powered by a recommender system that provides each user their own personalized tour recommendations, also keeps a database of user's preferences like

Preference Keywords:

Climate: Spring, Rainy, summer and winter.

Average expense: How much will the user be willing to pay for the Trip?

Preferred method of travel: By Flight, By Car, By Public Transports like Train or Bus, etc.

Average time: Duration for the trip

By using these keywords app recommends users to visit their personalized tour destinations.

The design of the backend is using NLP to get recommendation based on the reviews past users had during their stay in a hotel or travelling a tour. It uses topic modelling methods like NMF and LSA to extract features from a tour's past reviews. It also considers other data like tweets about the tour, climate, the best time to visit and average expenditure during the tour.

By using Twitter tweets and latest reviews in the recommender system ensures that the database is updated to current affairs about the tour as outdated data can provide false data about the tour.

C. Phase 3

This model provides business to local shop owners and restaurant through Xenia as they can use this platform to provide advertisement to users, this will help user to have better experience in during their tour, as they can use these advertisements to choose the best hotels or place to stay, and store will provide user to shop local products without spending time on shopping for local items and spend more time on tour.

This business model provides mutual benefits to both user and local shops and restaurants,

App collects data from local shop owners and restaurants like:

Questions (examples):

a. What are the best places to visit around their locations?

b. What is the best food to eat around their location?

c. What are the in demand items to shop around their location?

d. What precautions to take?

This type of crowd source data will help both the recommender system and the users to get more relevant information about the tour.

D. Natural Language Processing (NLP)

Natural language processing (NLP) is the crossing point of software engineering, semantics and AI. The field revolves around correspondence among algorithms and people in characteristic language and NLP is tied in with causing algorithms to comprehend and create human language. Utilization of NLP strategies incorporate voice controlled AI like Amazon's Alexa and Apple's Siri, yet in addition to things like machine interpretation and text-filtering

Following are Some NLP Techniques:

a) Parsing

What is parsing? As per the word reference, to parse is to "determine a sentence into its segment parts and depict their syntactic jobs."

That really nailed it, yet it very well may be somewhat more thorough. Parsing alludes to the conventional investigation of a sentence by an algorithm into its constituents, which brings about a parse tree showing their syntactic connection to each other in visual structure, which can be utilized for additional handling and comprehension.

b) Stemming

Stemming is a method that comes from morphology and data retrieval which is utilized in NLP for pre-preparing purposes.

Essentially, stemming is the way toward diminishing words to their promise stem. A "stem" is the piece of a word that stays after the evacuation of all joins. For instance, the stem for "contacted" is "contact." "Contact" is likewise the stem of "contacting", etc.

c) Text Segmentation

Text segmentation in NLP is the way toward changing content into significant units like words, sentences, various points, the hidden expectation and that's only the tip of the iceberg. For the most part, the content is sectioned into its segment words, which can be a troublesome assignment, contingent upon the language. This is again because of the intricacy of human language. For instance, it functions admirably in English to isolate words by spaces, aside from words like "icebox" that have a place together however are isolated by a space. The issue is that individuals in some cases additionally compose it as "ice-box."

E. Natural Language Processing (NLP)

Topic Modeling falls under unaided AI where the archives are handled to acquire the relative points. It is a vital idea of the customary Natural Processing Approach due to its capability to get semantic connection between words in the archive groups. Likewise, it has various different applications in NLP [9].

A portion of the notable ways to deal with perform theme demonstrating are

1. Non-Negative Matrix Factorization (NMF)
2. Latent Semantic Analysis (LSA)

a) Non-Negative Matrix Factorization (NMF)

Non-Negative Matrix Factorization is a measurable technique to decrease the element of the info corpora. It utilizes factor examination technique to give nearly less weightage to the words with less lucidness.

For an overall case, consider an information grid V of shape m x n. This strategy factorizes V into two lattices W and H, with the end goal that the element of W is m x k and that of H is n x k. The representation for the same is shown in Fig. 1. For the circumstance, V address the term archive grid, each line of lattice H is a word inserting and every section of the network W address the weightage of each word get in each sentence (semantic connection of words with each sentence). It is easy to track down a commonsense application with a model underneath.

However, given that all of W and H's entries are positive, the expectation is that all of W and H's entries are positive.

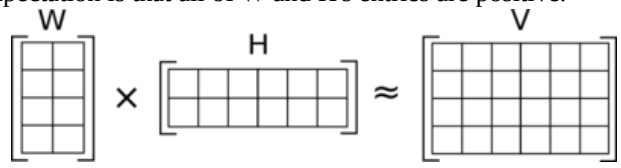

Fig. 1. Non Negative Matrix Factorization.

Presently, take a gander at the system for the situation. Assume this model has a dataset consisting of surveys of superhuman motion pictures. In the archive term lattice (input network), it has singular records along the lines of the grid and every interesting term along the segments. In the event that, the survey consists of writings like Tony Stark, Iron Man, Mark 42 among others. It very well might be gathered under the point Iron Man. In this technique, every one of the individual words in the report term grid are considered. While factorizing, every one of the words are given a weightage dependent on the semantic connection between the words. However, the one with most elevated weight is considered as the theme for a bunch of words. So this cycle is a weighted amount of various words present in the records.

I. Math behind NMF

As referenced before, NMF is a sort of unaided AI. The principle centre of solo learning is the measurement of distance between the components. The distance can be estimated by different strategies. Some of them are named as Generalized Kullback–Leibler divergence, frobenius standard.

1. Generalized Kullback–Leibler divergence

It is a factual measure which is utilized to evaluate how one appropriation is not quite the same as another. Closer the estimation of Kullback–Leibler uniqueness to nothing, the closeness of the relating words increments. At the end of the day, the divergence value is less.

Allow us to take a gander at the troublesome method of estimating Kullback–Leibler divergence. The Equation (1) for ascertaining the divergence is given by.

$$kl\_\mathrm{div}(x,y) = \begin{cases} xlog(x/y) - x + y & x > 0, y > 0 \\ y & x = 0, y \geq 0 \\ \infty & otherwise \end{cases} \quad (1)$$

2. Frobenius Norm

The other technique for performing NMF is by utilizing Frobenius norm. It is characterized by the square foundation of amount of supreme squares of its components. It is otherwise called eucledian standard. The equation (2) is given beneath.

$$\left||A|\right|_F \equiv \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\left|a_{ij}\right|^2} \quad (2)$$

A streamlining cycle is compulsory to improve the model and accomplish high exactness in discovering connection between the points. There are two sorts of improvement calculations present alongside scikit-learn package.

Coordinate Descent Solver

Multiplicative update Solver

NMF in action

For what reason would it be advisable for us to hard code everything without any preparation, when there is a simple way? Bundles are refreshed day by day for some demonstrated calculations and ideas. This model has a scikit-learn package to do NMF. It will utilize the 20 News Group dataset from scikit-learn datasets. Furthermore, it will initially import every one of the necessary packages.

Presently, this system will change over the record into a term-document matrix which is an assortment of the multitude of words in the given report.

Characterizing term document matrix is out of the extent of this article. To sum things up, the algorithm parts each term in the record and appoints weightage to each word.

a) Latent Semantic Analysis

The Latent Semantic Analysis model is a hypothesis for how meaning portrayals may be gained from experiencing enormous examples of language without unequivocal headings regarding how it is organized [10].

The issue at the hand isn't directed, that is this system don't have fixed marks or classifications appointed to the corpus.

To separate and comprehend designs from the archives, LSA naturally follows certain suppositions:

1) Meaning of Sentences or Documents is an amount of the importance of all words happening in it. Generally speaking, the importance of a specific word is a normal across every one of the archives it happens in.

2) LSA expects that the semantic relationship between words are available not unequivocally, however just idly in the huge example of language

Mathematical Perspective

Latent Semantic Analysis (LSA) involves certain numerical activity to get knowledge on a document. This calculation frames the premise of Topic Modelling. The centre thought is to take a matrix of documents and terms and decay it into a different document-topic matrix and a topic-term matrix. The initial step is creating an archive term grid utilizing Tf-IDF

Vectorizer. It can likewise be developed utilizing a Bag-of-Words Model, yet results are meagre and don't give any importance to the matter. Given m documents and n-words in the vocabulary, this model can develop an m × n matrix A where each line addresses a document and every segment addresses a word.

$$w_{i,j} = tf_{i,j} \times \log \frac{N}{df_i} \quad (3)$$

Equation (3) represents Term Frequency- Inverse Document Frequency (Tf-IDF) Score

Naturally, a term has a huge weight when it happens as often as possible across the report yet rarely across the corpus.

This study has structured a record term grid, A utilizing this transformation method (tf-IDF) to vectorise the corpus. (As it were, that the further model can measure or assess since it doesn't chip away at strings at this point!).

In any case, there is an unobtrusive disadvantage, this model can't construe anything by noticing A, since it's a noisy and sparse matrix. (At times too enormous to even consider evening figure for additional cycles). Since Topic Modelling is inherently an unsupervised algorithm, it needs to determine the latent topics in advance.

It is analogues to K-Means Clustering with the end goal that it determines the quantity of clusters beforehand. For this situation, it plays out a Low-Rank Approximation using a Dimensionality reduction technique using a Truncated Singular Value Decomposition (SVD) Singular value decomposition is a procedure in linear algebra that factorizes any matrix M into the result of 3 separate matrices: M=U*S*V, where S is a diagonal matrix of the singular values of M.

Truncated SVD decreases dimensionality by choosing just the t largest singular values, and just keeping first t columns of U and V. For this situation, t is a hyper parameter that this model can choose and acclimate to mirror the quantity of subjects it needs to discover.

$$A \approx U_t S_t V_t^T \quad (4)$$

Equation (4) represents Truncated SVD: U and V are Orthonormal Matrices, S being Diagonal Matrix. The Pictorial Representation is shown in Fig. 2.

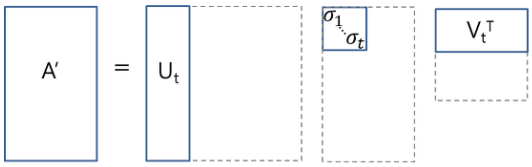

Fig. 2. Pictorial Representation of Truncated SVD.

With these document vectors and term vectors, this recommender system can now effectively apply measures, for example, cosine similarity to assess:

The similarity of various reports.

The similarity of various words.

The similarity of terms (or "inquiries") and documents (which gets valuable in data recovery, when need to recover entries generally pertinent to hunt question.

IV. RESULTS

This study aimed to develop the android application named Xenia. This app will help the tourist in a very effective way by recommending the smart options to them. Also, it will help the local businesses by providing them the advertisement in exchange for useful data for the recommender system.

Some Screenshots of the Xenia App is as follows:

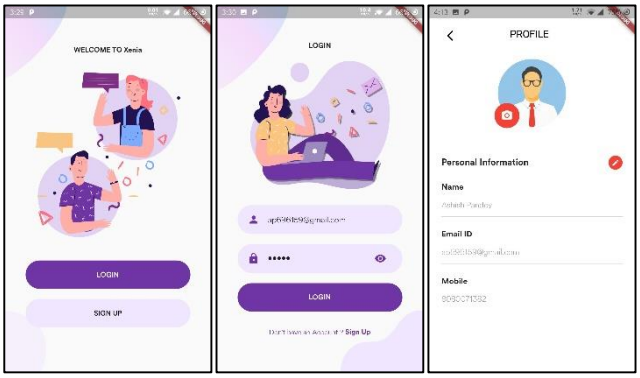

Fig. 3,Fig. 4, Fig. 5. Sign up Screen, Login Screen and Profile Screen.

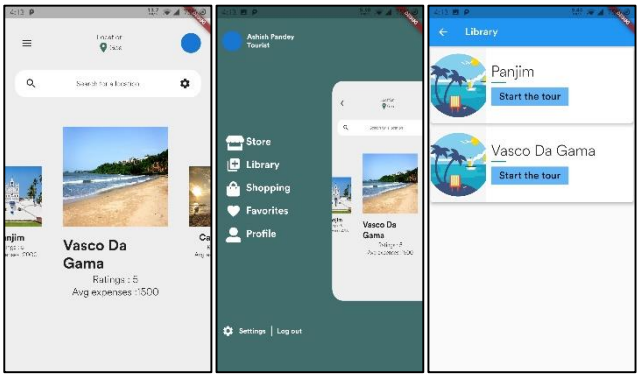

Fig .6, Fig. 7, Fig. 8. Home Screen, Menu Tab and Library Screen.

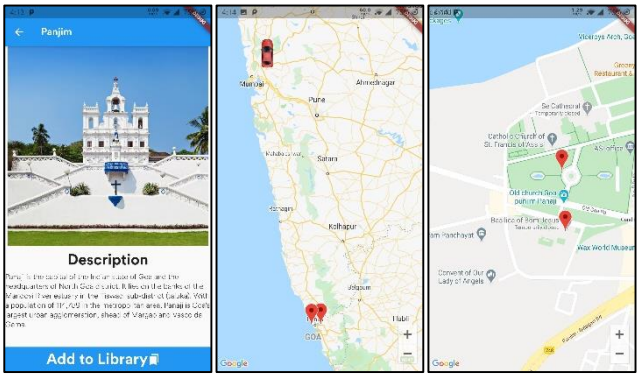

Fig. 9, Fig. 10, Fig. 11. Description Screen, Navigation Screen and Recommendation Screen.

V. CONCLUSION

This paper presented the design and implementation of a mobile application called Xenia, with which mobile users can get valuable recommendation on planning their trip. As shown in the result, this app has very informative UI which starts from login screen. The next screen will include store, by which user can shop the tour plans and add it in the library. Then it has Shopping and Favorites tabs. The screen shot of the maps and guides is also attached. By Xenia, users can get detailed information about important tour locations in text and pictures. In particular, this application can provide user location-based, review based and climate based recommendation for tour planning. This app also provides a business model for local stores, hotels and restaurants in exchange for the genuine data of that place. This project will also help in growing the local businesses and the valuable data side by side using

crowdsourcing. The recommender system will be getting day by day as its data is totally gathered from locals of that place itself. The ultimate goal of this project is to give the user a one-of-a-kind informative and interactive experience and allow them to have all features they desire in the palm of their hand by covering every possible approach of providing information and recommendations.

REFERENCES

[1] P.Mate and H. Chavan, "Android Based Continuous Query Processing in Location Based Services", International Journal of Advanced Research in Computer Science and Software Engineering, 2014, Vol-4, Issue-5.

[2] S.Bhattacharya and M. B. Panbu, "Design and Development of Mobile Campus, an Android based Mobile Application for University Campus Tour Guide", International Journal of Innovative Technology and Exploring Engineering, 2013, Vol-2, Issue-3.

[3] Brown, B. and M. Chalmers, "Tourism and Mobile Technology", Proceedings of Eighth European Conference on Computer-Supported Cooperative Work, 2003.

[4] U. Thakur, V. Nagare and M.Shimpi, "Tour Guide System using Augmented Reality: A Review", International Journal of Soft Computing and Engineering (IJSCE),2015, Vol-5 Issue-5.

[5] J.Choi B.Jang, G..J.Kim, "Organizing and presenting geospatial tags in location-based augmented reality". Verlag London Limited, 2010.

[6] R. A. Abbaspour* and F. Samadzadegan, "Building a Context- Aware mobile tourist Guide system based on service oriented architecture", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2008, Vol-37, Issue-4.

[7] Dadape Jinendra R., Jadhav Bhagyashri R., Gaidhani Pranav Y., Vyavahare Seema U., AchaliyaParag N., "Smart Travel Guide: Application for Android Mobile", 1st International Conference on Recent Trends in Engineering & Technology, 2012, Special Issue of International Journal of electronics, Communication & Soft Computing Science & Engineering, ISSN: 2277-9477.

[8] M. Kenteris, D. Gavalas and D. Economou, "An innovative mobile electronic tourist guide application", Personal Ubiquitous Comput, 2009, Vol-3, Issue-103.

[9] Aurélien Géron, "Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems 1st Edition".

[10] Benjamin Bengfort, Rebecca Bilbro, Tony Ojeda "Applied Text Analysis with Python: Enabling Language-Aware Data Products with Machine Learning".