Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

An Analysis of Location-Based Information for Exploring Geo-locational Data

R Vijaya Saraswathi
Department of CSE, VNRVJIET, Hyderabad.

***Abstract*— This research introduces a novel geolocational clustering approach tailored to address the intricate task of selecting an optimal residence location in Hyderabad, India. The process of choosing a suitable place to live, often faced by students and job seekers, is fraught with complexity and uncertainty. To simplify this decision-making process, we harness geolocational data analysis, incorporating variables such as food preferences, exercise routines, and other lifestyle factors. The methodology employs the K-means clustering algorithm to create clusters on the map, with the optimal number of clusters determined through the elbow method. To enhance the accuracy and robustness of the model, we explore alternative clustering methods, including K-modes, K-medoids, and affinity propagation. Beyond its applicability to students and job seekers, our approach holds broader relevance, offering valuable tools for commercial enterprises and tourist services. It automates the analysis of diverse customer preferences, thereby reducing the workload for tourist guides. As geolocational data analysis gains prominence, its potential applications in optimizing various aspects of daily life become increasingly evident, streamlining routine tasks and enhancing overall experiences. In summary, our geolocational clustering method serves as a powerful resource for informed location selection in Hyderabad, promising streamlined decision-making processes and improved quality of life across diverse domains.**

***Keywords*— Exploratory Analysis, K-means clustering, K-medoids clustering, k-modes clustering, Affinity Propagation, Geo-Location.**

I. INTRODUCTION

The research paper inquiries into the potential for using ML(ML) algorithms for exploratory analysis of geolocational data to recommend locations. Using the existing research on clustering techniques of ML, predictive models, and data visualizations, the Main goal is to determine whether ML can be used to accurately predict the suitable locations for a variety of potential activities. Through the use of available datasets, the research will provide an overview of the current geolocational data that is already being used to recommend locations and how ML can be used to effectively supplement or improve upon this analysis. The outcomes derived from experiments conducted on geolocational data will serve as a means to comprehensively evaluate the efficiency of ML algorithms within this domain. This research Endeavor aims to provide a deeper understanding of the applicability of ML techniques in the domain of location recommendation and to discern their suitability across diverse scenarios. To generate this model many attempts are made to identify the major attributes of a data set and further it is investigated to generate different interpretations, attention is focused on model specification and parameter estimation. This generation of the model involves the use of K-Means Clustering, K- modes, k-medoids, and Affinity Propagation to find the best accommodation for people who are searching for a location to live in any city of their choice by classifying accommodation for people who are new to the place based on their preferences on services and closeness to a location. For an individual like "A" who has recently relocated to a specific area, making an informed decision about where to reside is challenging, given their existing preferences and inclinations, illustrated in Figure 1. This ML-based model offers valuable assistance by considering all of the user's preferences to recommend a suitable location. It is particularly beneficial for newcomers, such as students seeking hostels, entrepreneurs looking to establish businesses, and tourist guides aiming to connect with people interested in specific destinations. The model utilizes clustering, where the user is assigned to the cluster that aligns best with their interests, preferences, and needs, based on distance metrics. This places the user in the vicinity of the most fitting location among available options, conserving valuable resources. for the user and making their life simpler, as an individual(A) would be in a place of their own interest as everything, they prefer is nearby to the cluster they are placed in. It would save time and effort of search for both the student and the food providers. Students moving to new place can conveniently search for a place as in Figure 2. Convenience means better sales and saving time for the customer.

This kind of analysis is also useful for business magnets which includes hotel owners, restaurant owners, and gym owners as they can know the preferences and tastes of customers and expand their business in accordance with the requirement of the customers. This also is useful for those who want to start a new business as this would allow the person to know the requirements of the customers living nearby saving a lot of effort for the person who wants to start a business and maximize their profits. A similar kind of problem is being faced by employees as well who move to a new place in search of a job. This would make it easier for you to choose a place to stay and lead a comfortable life. The analysis of geolocational data has been useful for many kinds of applications as any location recommendation engine or system is built on top of the exploratory analysis of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

geolocational data. A ML model is prepared by exploratory analysis of already existing Geolocational data using clustering algorithms, which recommends a location based on their preferences through input.

Geolocated tweets are not uniformly distributed across geographical space; instead, they tend to cluster in specific areas. This phenomenon is evident through an examination of a three-month dataset of geolocated tweets within London, this work analyses tweet hotspots and demographic characteristics of the wards where these hotspots appear [2]. In this research mobility challenges are reviewed, from a number of colorful outlooks, containing statistical study, dossier junction, pattern finding and connected dossier. To be specific, we present a primary dossier analysis accompanying Civilians for public conveyance service toward a record of what happed in Curitiba, Brazil [3]. A decision support system (DSS) involving geolocational data for analysis and evaluation for different kinds of transport is being shown [4]. These provide a base for us to understand how an analysis on geolocational data would help in building complex application that would solve greater problems of man-kind. In this, there are multiple algorithms that are used which support one another and affinity propagation backs the claims made by clustering algorithms.

A. Advantages of Geolocational Analysis

Geolocational analysis is used for analyzing different geolocational spaces. It helps in taking decisions, taking customer insights, managing risks. Geolocational analysis can provide insights on customer behavior and preferences, which helps in making informed decisions on resource allocation and expansion. This also helps in identifying natural disasters and other risks, which helps in measuring or minimizing damage or losses. Thus, geo-locational analysis is an important tool for businesses requirements to reduce their operations and improve their business.

Figure 1. Representation of Various Facilities Around a Specific Location.

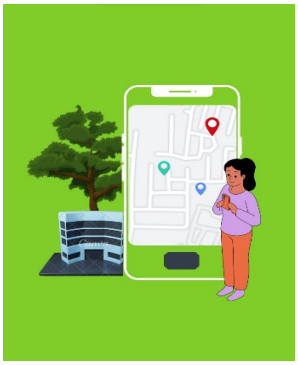

Figure 2. A Student Searching for a Location According to his/her Preferences.

II. RELATED WORK

Location and GIS should encourage the use of modernizing approaches that take into account spatial variability and spatial autocorrelation from a methodology standpoint. These modernization approaches may offer many practical alternatives to conventional competition analysis in the lodging sector (Nicholls, 2019). It is possible to collect the geographic data on the Internet utilizing system that enables constant updating of each establishment's information and, as a result, monitoring of the level of competition in a destination, is illustrated in one section, examine diabetes-related participation on Twitter by describing the frequency and timing of diabetes-related tweets, the geography of tweets, and the types of participants over a 2-year sample of 10% of all tweets [1].

Geolocated tweets are not evenly spread across space, but appear in accumulations. By exploring a collection of 3 months of geolocated tweets for London, this work analyses tweet hotspots and demographic characteristics of the wards where these hotspots appear [2]. In this research mobility challenges are reviewed, from a number of colorful outlooks, containing statistical study, dossier junction, pattern finding and connected dossier. To be specific, we present a primary dossier analysis accompanying Civilians for public conveyance service toward a record of what happed in Curitiba, Brazil [3]. A decision support system (DSS) involving geolocational data for analysis and evaluation for different kinds of transport is being shown [4]. Unborn Mobility seeing is an innovative smartphone- grounded trip check that was field- tested in previous work, together with the Household Interview trip check (successes), in Singapore. This paper presents the findings of an exploratory analysis of the data collected in this test. The clustering of the day patterns from the FMs (Foundational Models) data revealed a large day- to- day variability of stoner that couldn't be captured by a shot with a 1- day check [5]. The results show that monumental trees are spatially concentrated in high-income neighborhoods, and this fact represents an index of environmental inequality. This opinion can give support for decision- making [6]. The model used in this research helps in

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

identifying patterns of road networks using characteristics of the network and demands of business. The consumption of the energy and emigration of contaminant computation whose methodology was developed by working group of company [7]. This design uses the K- Means clustering system to find the emigrant harbors by ranking migratory, harbors according to their amenities, budget and propinquity preferences. Cost, cleans, assay and summations K- Means on geolocation data to recommend indigenous accommodation in the megacity [8]. The main aim is to classify Parkinson's based on voice attributes in speech using machine state algorithms [9]. The goal is to classify the different images of the Parkinson's diseases with the help of ML algorithms [11]. Designing Geodatabases for Transportation addresses the development of Civilians to manage data relating to the transportation installations and service generally organized around the modes of trip for accurate and dependable data exchange. Transportation involves several modes of trip, and although the details of each mode can be relatively different, this book demonstrates how all follow an introductory abstract structure [12].

The goal of the application is to guide the scholars and keep them streamlined about what is passing in the class. Using mobile apps scholars learn from any geographical position but have also helped them use the technology at their disposal in judicious ways [13]. Using camera of laptop, we capture the images and detect the eyeball movement by indexing of face by using neural networks and open cv and detect the physical movements of a mouse [14]. The rest of the paper is organized as follows: Section 3 discusses the Proposed Work, Section 4 elaborates on the Results and Discussion, followed by Conclusions.

III. PROPOSED METHODOLOGY

Figure 4 describes the working of the proposed model. The diagram shows various phases such as data collection, model construction, and clustering process etc. All the steps involved in the flow chart are described below and measures are taken to enhance the results. Whereas, Figure 3 shows the feature selection process that enables to select optimal features on which clustering is applied.

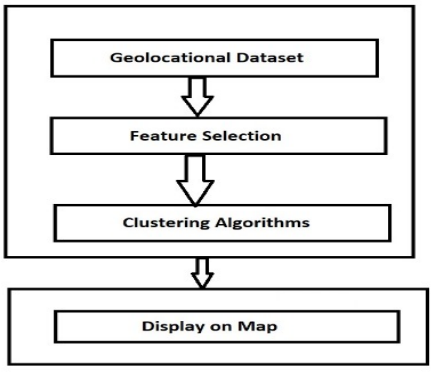

Figure 3. The System Architecture

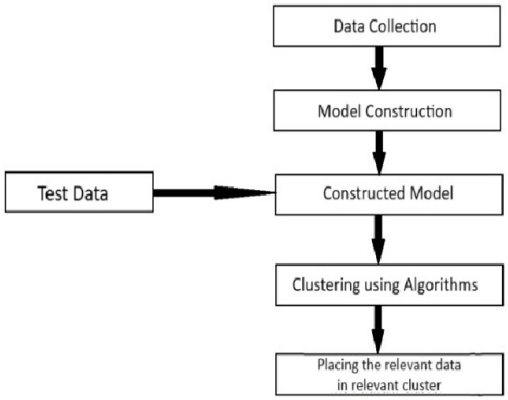

Figure 4. The General flow Diagram of Proposed Work

In Figure 5, the input data is divided into training and testing data set. Then the training data is fed into k-means clustering, k-medoids clustering, k-modes clustering and Affinity Propagation Algorithms (clustering algorithm that identifies exemplars in a dataset to form clusters based on similarity measures) that to obtain a models. The ML process is as follow:

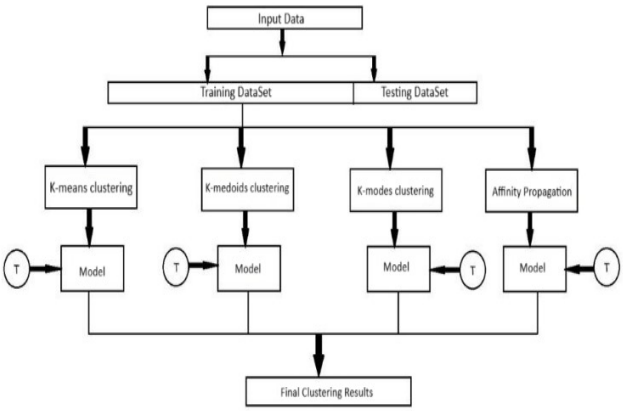

Figure 5. The Flow Diagram of Proposed Methodology

A. Data Cleaning

The dataset selected is taken form Kaggle, named as food coded [16] and it contains 125 rows and 61 columns. The features selected for the model did not contain some null values and the tuples were removed from consideration.

B. Data Preparation

The feature selection step was done manually by taking into consideration some of the important criterions which have a significant effect on selecting a place to stay. These features encompass an array of attributes, ranging from cooking habits, dining preferences, employment status, food preferences, exercise routines, income levels, to interest in sports, and more. The null values from the dataset were dropped. The

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

SNS-pair plot was used to see how many different values were present in the dataset and their count. The boxplot method was used to determine the relationship between two columns as shown in Fig-5. The manual interpretation of the interrelationships among the columns is used as the basis for selecting the columns in the clusters.

C. K-means Clustering

The initial step of selecting the number of clusters is executed with the help of elbow method (The elbow method is a technique for determining the optimal number of clusters in a dataset by identifying the "elbow point" on a plot of within-cluster sum of squares against the number of clusters). As Shown in the Fig-7 the number of clusters that would be optimal based on the elbow method was 6 and the number of clusters selected were 6. Then the K-means clustering was done with the help of SKLearn package.

Thus, the value of k-selected is 6.

1. K-centroids were randomly selected from the dataset.
2. Each datapoint is assigned to their closest centroid forming the predefined k- clusters.
3. The variance of each cluster is calculated and new a centroid is selected
4. The above 3 steps were repeated until there is no new reassignment.
5. The model with 7 clusters is formed.

$$similarity = \cos(\theta) = \frac{xi.yi}{||xi||||yi||} = \frac{\sum_{i=1}^{n} xi.yi}{\sqrt{\sum_{i=1}^{n} xi^2}\sqrt{\sum_{i=1}^{n} yi^2}} \text{-----(1)}$$

The similarity as shown above is calculated and clusters are formed.

$$Mean = \frac{\sum of\ Data\ Points}{Total\ no.of\ Data\ points} \text{----------(2)}$$

is used to recalculate the centroids. Algorithmic approach for K-modes Clustering

It is the similar step as followed from the K-means clustering method. We have found the number of clusters that can be formed using the elbow method. As showed in the Fig-8 the number of clusters that would be optimal based on the elbow method was 5 and the number of clusters selected were 5. Then the K-modes clustering was done with the help of sklearn package.

Thus, the value of k selected is 5.

1. Assigned each data point to the closest centroid based on the similarity metric used. The similarity metric is based on matching categorical attributes.
2. For each cluster, the mode of the categorical attributes among the assigned data points is computed. The mode Represents the most common value of the attribute within the cluster.
3. The centroids are updated based on the new modes.
4. The above 3 steps are repeated until convergence.

The distance among the data objects in the dataset is calculated by the similarity and dissimilarity measurement functions. Here in the scenario of K-modes, these distances are computed by a dissimilarity metric function also known as hamming distance, which is the no. of categorical attributes that change among the two data objects.
Let p and q be two categorical data objects that are defined by z features or attributes.

$$d(p,q) = \sum_{j=1}^{z} \delta(p_j, q_j) \text{----------(3)}$$

Where $\delta(p_i, q_i) = \begin{cases} 0 \ if \ p_j = q_j \\ 1 \ if \ p_j \neq q_j \end{cases}$

Mode = max (all occurrences in the cluster) is used to recalculate the centroids.

D. K-medoids Clustering

It is the similar step as followed from the K-means clustering method. We have found the number of clusters that can be formed using the elbow method. As showed in the Figure 9 the number of clusters that would be optimal based on the elbow method was 7 and the number of clusters selected were 7. Then the K-modes clustering was done with the help of sklearn package.

Thus, the value of 'k' selected was 7.

1. Based on a distance metric the datapoint is assigned to the closest medoid.
2. A new medoid is selected for each cluster that minimizes the sum of distances between the medoid and the other data points in the cluster. This was done by trying out candidate medoids in each cluster.
3. The medoids of the cluster are replaced by newly selected medoids.
4. The above 3 steps are repeated until convergence (Convergence in machine learning refers to the point at which an algorithm has reached a stable and optimal solution, typically indicated by minimal or no further improvement in a performance metric).

A medoid is the point in the cluster with the least dissimilarity to all other points within a cluster. The dissimilarity that lies among the medoid (Ai) and the object (Bi) can be computed using E = |Bi - Ai|.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

$$c = \sum_{ci} \sum_{pieCi} |Pi - Ci| \quad \text{----------(4)}$$

Similarity Index is used to calculate the distance between 2 tuples in the dataset and median is used to recalculate centroids.

E. Rest API used for Map

We have used The Folium Rest API for the visualization of the particular place that is considered for the project, like Hyderabad in this case. This Rest API is used for pointing out the clusters formed on the map and it also helps in differentiating the clusters formed with different colors. Figure 10 displays an example Folium map.

F. Affinity Propagation

We have used Affinity Propagation for specifying the number of clusters formed as it is quite convenient when you cannot specify the number of clusters, and it's suited for geospatial data as it works well with non-flat geometry.

Algorithm 1:

1: Input: Similarity matrix (S), maximum number of iterations (max_iter)
2: P = number of data points
3: K = number of clusters
4: Initialize availability matrix (A) and responsibility matrix (R) to zeros
5: for j = 1 to max_iter
6: # Update responsibility matrix (responsibility updates)
7: for p = 1 to P:
8: for k = 1 to K:
9: # Compute responsibility of point p to cluster k
10: temp_sum = S[p][k] + max(a[p'][k'] for k' ≠ k)
11: R[p][k] = S[p][k] - temp_sum
12: # Update availability matrix (availability updates)
13: for p = 1 to P:
14: for k = 1 to K:
15: if p ≠ k:
16: # Compute availability of point p from cluster k
17: temp_sum = max(0, R[k'][k]) for k' ≠ k
18: A[p][k] = min(0, R[p][p] + temp_sum)
19: # Update cluster assignments
20: for p = 1 to P:
21: # Assign data point p to the cluster with the maximum (A[p][k] + R[p][k])
22: cluster_assignment[p] = argmax_k (A[p][k] + R[p][k])
23: # Check for convergence (no change in cluster assignments)
24: if cluster_assignment does not change from previous_iteration_assignment:
25: break
26: # Update cluster exemplars (centroids)
27: for k = 1 to K:
28: # Find data point that has the maximum (A[p][k] + R[p][k])
29: cluster_exemplar[k] = argmax_p (A[p][k] + R[p][k])
30: # Output: Optimal centroids (cluster_exemplar) and cluster assignments (cluster_assignment)

IV. RESULTS AND DISCUSSION

Exploratory data analysis (EDA) is an approach of analyzing data sets to summarize their main characteristics, often using statistical graphics and other data visualization methods. This generation of the model involves the use of K-Means Clustering, K-modes, k-medoids, and Affinity Propagation to find the best accommodation for people who are searching for a location to live in any city of their choice by classifying accommodation for people who are new to the place on the basis of their preferences on services and closeness to a location. This benefits not only students but also assists restaurant managers in the growth of their establishments.

This work also involves recommending hotels, gyms and other needs for the user who has accommodated to an area newly. It is difficult for a user to find all the places in a newly accommodated area. So, it is easy if we recommend nearby places. The existing system is either the conventional visiting the different places nearby your desired location or we must go to the desired location and ask for the Google Assistant to suggest the place to be visited nearby the location. In either case we personally have to make the trip to location to visit the desired place. It is a hectic process to go and visit the location and check out each and every place over there, sometimes we may not get satisfied with the result. But by forming Clusters based on their preferences it becomes easy for user. These clusters can be used by users for selecting a place to stay. This analysis as of now is done only on Hyderabad and also place recommendation system according to their needs, Preferences and budget is not available. So, we can further improvise by adding recommendation system which not only analyses the places according to the preferences but when given the preferences of the user also suggest the preferred place for him to accommodate. This analysis can also be expanded beyond Hyderabad on the map according to the user's requirements.

The SNS-pair plot, as shown in Figure 6, is used which will allow to recognize the distinct classes in each column and the number of tuples falling in each of the category which provides a clear picture of dataset.

The boxplot as shown in Figure 7 consists of distribution of dataset through their quartiles. The median, lower and upper quartiles mentioned provided us with a better understanding of data and a quick and easy way of feature selection. The elbow method as shown in Figure 8 helped us to identify the optimal number of clusters to be created in the clustering algorithm.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

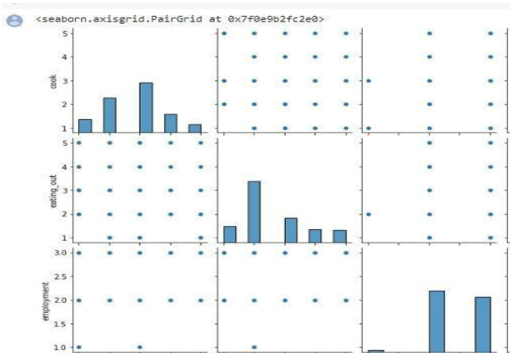

Figure 6. Output of SNS-pair Plot for 3 Columns.

It involves pointing the with in cluster sum of squares and the number of clusters in the algorithm and selecting the number of clusters at the elbow of the plot

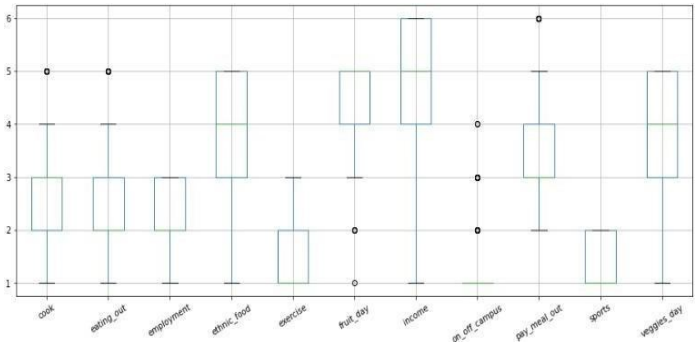

Figure 7. Output of Boxplot Method on Selected Features

The intuition is that increasing the number of clusters will naturally improve the fit (explain more of the variation), since there are more parameters (more clusters) to use, but that at some point this is over-fitting, and the elbow reflects this.

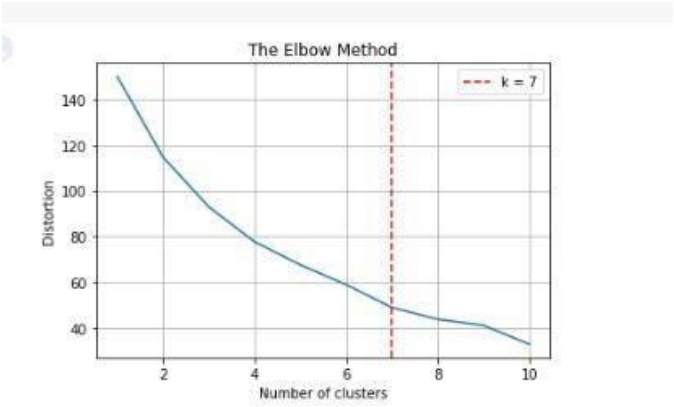

Figure 8. The Elbow Method with Red Line Showing Number of Clusters to be Considered.

For example, given data that actually consist of *k* labelled groups, *k* points sampled with noise – clustering with more than *k* clusters will "explain" more of the variation (since it can use smaller, tighter clusters), but this is over-fitting, since it is subdividing the labelled groups into multiple clusters.

The idea is that the first clusters will add much information (explain a lot of variation), since the data actually consist of that many groups (so these clusters are necessary), but once the number of clusters exceeds the actual number of groups in the data, the added information will drop sharply, because it is just subdividing the actual groups. Assuming this happens, there will be a sharp elbow in the graph of explained variation versus clusters: increasing rapidly up to *k* (under-fitting region), and then increasing slowly after *k* (over-fitting region).

```
Best run was number 1
Best run was number 1
Best run was number 5
Best run was number 4
Best run was number 4
Best run was number 3
Best run was number 3
Best run was number 4
Best run was number 5
Best run was number 1
```

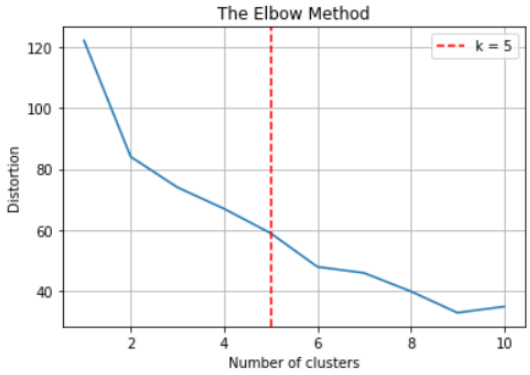

Figure 9. The Elbow Method with Red Line Showing Number of Clusters to be Considered.

The elbow method as shown in Figure 10 helped us to identify the optimal number of clusters to be created in the clustering algorithm. It involves pointing the with in cluster sum of squares and the number of clusters in the algorithm and selecting the number of clusters at the elbow of the plot.

The elbow method helped us to identify the optimal number of clusters to be created in the clustering algorithm. It involves pointing with in cluster sum of squares and the number of clusters in the algorithm and selecting the number of clusters at the elbow of the plot. The folium image as shown in Figure 11 provides an easy and flexible way to explore and use, explore and visualize geolocational data.

Affinity algorithm is used to group similar items into clusters based on their pairwise similarity which is shown in Figure 12. The main advantage of affinity propagation is that it does not require a predetermined number of clusters, unlike other clustering algorithms such as k-means or hierarchical clustering. Instead, it automatically determines the optimal number of clusters based on the data.

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

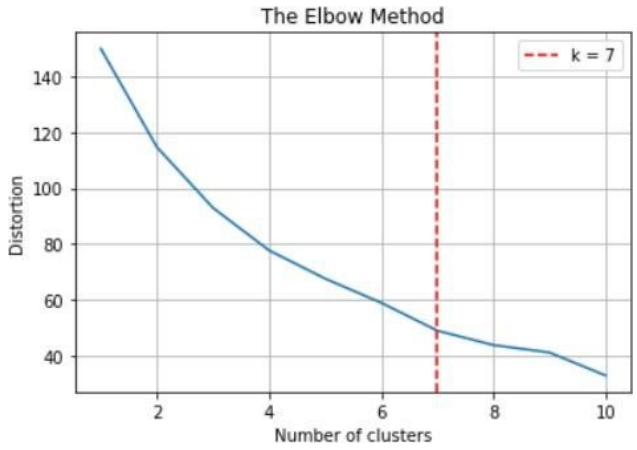

Figure 10 The Elbow Method with Red Line Showing Number of Clusters to be Considered.

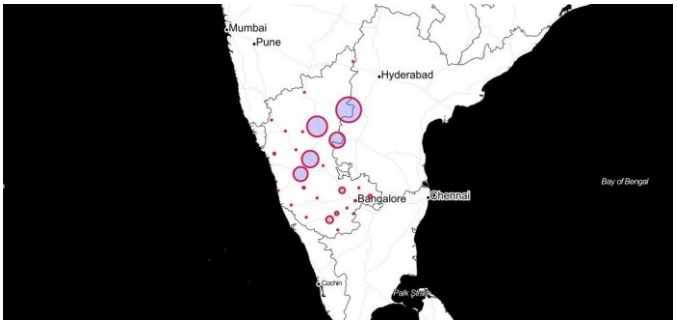

Figure 11. Sample Folium Image

```
[ ]  model = cluster.AffinityPropagation()

     k = n["cluster"].nunique()
     sns.scatterplot(x="lat", y="long", data=n,
                     palette=sns.color_palette("bright",k),
                     hue='cluster', size="centroids", size_order=[1,0],
                     legend="brief").set_title('Clustering (k='+str(k)+')')

     Text(0.5, 1.0, 'Clustering (k=6)')
```

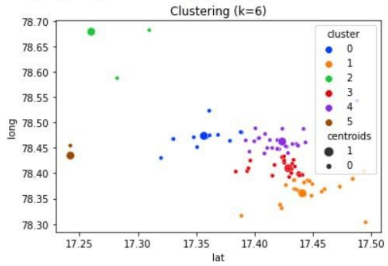

Figure 12. Implementation and Output of Affinity Propagation

The clusters generated, refer Figure 13, are labelled by using inverted drop shaped icon with a number on it. The mean or the centroid of cluster are marked by using the inverted drop shaped icon and the cluster points are located around the centroid.

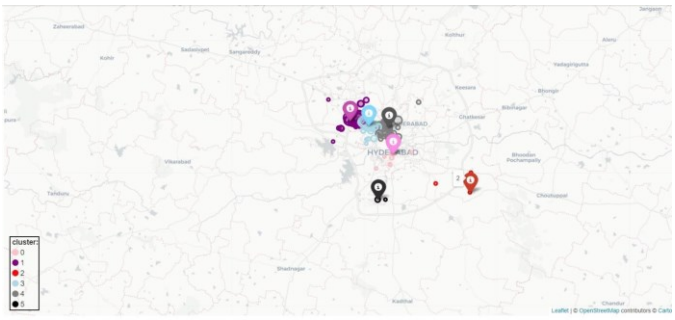

Figure 13. Clusters on the Map tells Different Preferences. The pointers (Inverted drop shaped Icon) on the map are cluster centroids.

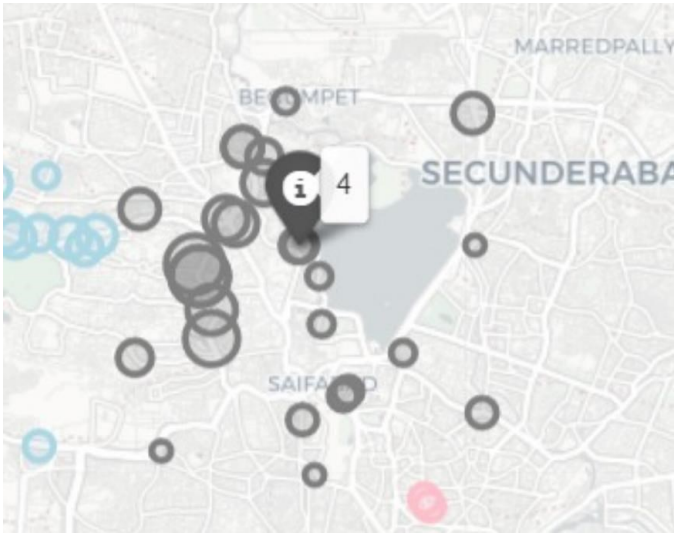

Figure 14. Points Around a Single Cluster

The cluster as an individual looks as shown in Figure 14 the centroid is represented using the inverted drop shaped icon and all the black points around it shows the datapoints that are close enough to fall under that cluster.

V. CONCLUSIONS

This analysis is easy to use because it is user-friendly and budget-friendly. A common problem of the migrant people is solved through this analysis. This analysis is used to find accommodation easily and its fits in our budget and it will more useful students who are studying. It not only useful for the students but also is very useful to the restaurant owners or the business owner who want to set up their new restaurant according to the choices or menu and the people preferences. Even this analysis makes the task easy for tourists guides, it can be a reference to the guide for tours about the tourists and accommodate them according to their preferences. This analysis makes the task of finding preferences according to their requirements easier, faster and comfortable to the user.

These days it become vital for people to live in a place where they could find their comforts in the proximity due to their busy life schedule and other factors. This project helps them by giving a location to stay which matches their comfort. The same can be used business magnets to establish as well as flourish their business based on customers interest and preferences. In this way this analysis helps the migrants to easily find accommodation according to their preferences and budget.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 13 Issue 07, July-2024

REFERENCES

[1] Use of Social Media in the Diabetes Community: An Exploratory Analysis of Diabetes-Related Tweets. Yang Liu 1, PhD ; Qiaozhu Mei 1, 2 , PhD ; David A Hanauer 1, 3 , MS, MD ; Kai Zheng 1, 4 , PhD ; Joyce M Lee 5, 6 , MD, MPH , JMIR publications.

[2] Demography of Twitter Users in the City of London: An Exploratory Spatial Data Analysis Approach,Barbara Hofer, Thomas J. Lampoltshammer & Mariana Belgiu , Springer.

[3] Exploratory Analysis of Public Transportation Data in Curitiba,Nádia P. Kozievitch,Tatiana M. C. Gadda,Keiko V. O. Fonseca,Marcelo O. Rosa,Luiz C. Gomes Jr.,Monika Abkar,IEEE.

[4] Arampatzis, G., Kiranoudis, C., Scaloubacas, P., and Assimacopoulos, D. (2004). A gis-based decision support system for planning urban transportation policies. European Journal of Operational Research, 152(2):465 – 475. New Technologies in Transportation Systems.

[5] Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore, Fang Zhao,Francisco Câmara Pereira Rudi Ball, Youngsung Kim, Yafei Han, Christopher Zegras, Moshe Ben-Akiva, Sage journals.

[6] Approach to Urban Environmental Justice Using Exploratory Spatial Data Analysis. The Case of Valencia's Monumental Trees, Alfonso Gallego-Valadés,Francisco Ródenas-Rigla , andJorge Garcés-Ferrer, MPDI.

[7] Arampatzis, G., Kiranoudis, C., Scaloubacas, P., and Assimacopoulos, D. (2004). A gis-based decision support system for planning urban transportation policies. European Journal of Operational Research, 152(2):465 – 475. New Technologies in Transportation Systems.Barczyszyn, G. L. (2015).

[8] B.Buvanswari and T.Kalpalatha Reddy, "A Review of EEG Based Human Facial Expression Recognition Systems in Cognitive Sciences" International Conference on Enenrgy, Communication,Data analytics and SoftComputing(ICECDS),CFP17M55-PRJ:978-1-5386-1886-8",August 2017.

[9] Sharanyaa, S., P. N. Renjith, and K. Ramesh. "Classification of Parkinson's disease using speech attributes with parametric and nonparametric ML techniques." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2020.

[10] B.Buvaneswari and Dr.T. Kalpalatha Reddy,"EEG signal classification using soft computing techniques for brain disease diagnosis",Journal of International Pharmaceutical Research ,ISSN : 1674-0440,Vol.46,No.1,Pp.525-528,2019.

[11] Sharanyaa, S., P. N. Renjith, and K. Ramesh. "An Exploration on Feature Extraction and Classification Techniques for Dysphonic Speech Disorder in Parkinson's Disease." In Inventive Communication and Computational Technologies, pp. 33-48. Springer, Singapore, 2022.

[12] Butler, J. A. (2008). Designing Geodatabases for Transportation. Esri Press.

[13] Sharanyaa, S., and M. Shubin Aldo. "Explore places you travel using Android." In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 4796-4799. IEEE, 2016.

[14] Sharanyaa, S., and Madhumitha RP. "Eyeball Cursor Movement Detection Using Deep Learning." RP, Madhumitha and Rani. B, Yamuna, Eyeball Cursor Movement Detection Using Deep Learning (July 12, 2021) (2021).

[15] B.Buvaneswari and Dr.T. KalpalathaReddy,"ELSA- A Novel Technique to Predict Parkinson's Disease in BioFacial",International Journal of Advanced Trends in Computer Science and Engineering, ISSN 2278- 3091,Vol.8,No.1,Pp. 12-17,2019

[16] DatasetLink:- https://www.kaggle.com/datasets/borapajo/food-choices?select=food_coded.csv, Source: Kaggle.