

# **Food for the Needy**

## **PROJECT REPORT**

Submitted in fulfilment for the J-Component of Technical Answers for Real World

Problems (TARP)- ITE3999

**B. Tech – Information Technology**

By

18BIT0483 ANAMIKA GUHA

Under the guidance of

**Dr. ASWANI KUMAR CHERUKURI**

**SITE**



**VIT<sup>®</sup>**  
**Vellore Institute of Technology**  
(Deemed to be University under section 3 of UGC Act, 1956)

**School of Information Technology and Engineering**

**FALL Semester 2021-22**



## **PROBLEM STATEMENT**

In the real world out there, There is no common source for the people who are in need for food and the people who are willing to provide food for free. The sources available are just advertisements and manual information. To tackle this situation, we need a common source which is both user-friendly and effectively establishes this connection.

## **OBJECTIVES**

- To develop an android app, that will help NGOs to supply free food.
- To provide proper communication between NGOs and the people in need.
- Finding food in demand locations.
- Reduce the amount of food going into waste.
- Redistribute the food to the people who can't afford it and need it.

## **OUTCOME**

We developed an android app that eradicates the problem of connecting NGOs and the people in need. We can effectively communicate the right locations of the food requests and optimal locations of food stalls to NGOs such that they can manage the food to be provided at each stall. This makes sure that no food is wasted and every person is well fed.

## **RELEVANCE (TO WHOM THE PROBLEM & SOLUTION ARE RELEVANT)**

**To which section of society this problem is relevant?**

The people below poverty line , orphans , undernourished people who need food.

**How is the solution going to impact them?**

The solution can provide better food nutrition and nourishment to the people in need, in our society.

**Write in terms of social, economic, and technical impact you are going to make with the solution?**

To make people in the sections below poverty line well nourished, as good health ensures the decline of the death rate and a good economy.

Providing better services using modern technologies.

**Also how are you going to measure it??**

The number of hungry people whom we feed daily.

The number of people who are donating food making use of our services.

## **REAL WORLD PROBLEM AIMED TO SOLVE**

Though there is enough food in India, a lot of people in slums are suffering and even dying from hunger and starvation. This is mainly due to less knowledge and lack of proper communication with NGOs who are ready to distribute food for the people in need. This negatively impacts the economy of India. Currently, the NGOs willing to help the poor are getting information regarding people in need through their connections, identifying places by themselves by going around the city, etc. But this method of locating places where there is demand for food is quite inefficient. Also, once they find the places, they have to transport the food to the location from their main branches, which are sometimes located extremely far away. A new solution should be formulated such that the people in need can directly contact the NGOs by raising their voice and the locations which have more demand for food can be identified such that NGOs can have their food stalls installed at these locations for better service.

## **SOLUTION**

The application will use the Google Maps API and machine learning concepts to connect and locate people when they authenticate themselves in the application. Whenever they are signing up, the location will be saved and they will be directed to the nearest food stall.

The whole project is divided into two parts, the first part consists of the android app which will serve the people for finding the nearest food stalls to them, whenever they check-in, their location will be saved and they will be directed to the food stall. In case if there are no food stalls near their location, we will collect the data of the request and use the k-means algorithm to make the prediction, and provide the information to NGOs to find the most suitable location for them to set up a food stall and reach maximum people. Registrations will be saved in the database firebase and the application will be refreshed within regular intervals of time so as to find the new store locations or change the previous location so as to reach out and help the maximum people possible. And the second part consists of finding the optimal location of the food stalls so that they can be easily accessible to a large number of people. For the first part, we will be using Native Android Development Kit using Java as our primary language and Firebase for the database. And for the second part, we will be using a Machine Learning Algorithm called K-Means Clustering Algorithm to cluster groups of near people together.

## **WHAT IS THE INNOVATION & NOVELTY**

- Novel way of communication between NGOs and needy people
- Discovering and identifying food in demand locations

## **TECHNOLOGY STACK TO BE USED:**

K-Means Algorithm, Elbow Method, Android Studio(Google Maps API, Firebase Realtime DB, Firebase Authentication)

### **K-Means Algorithm:**

The algorithm takes the unlabelled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim of this algorithm is to minimize the sum of distances between the data point and their corresponding clusters.

Here, the dataset contains the x-coordinate and y-coordinate of the location from where the food stall requests have come and predicts the output which will be considered as the optimal location for the food stall.

### **Elbow Method:**

A fundamental step for any unsupervised algorithm is to determine the optimal number of clusters into which the data may be clustered. The Elbow Method is one of the most popular methods to determine this optimal value of k. To determine the optimal number of clusters, we have to select the value of k at the “elbow” that is the point after which the distortion/inertia start decreasing in a linear fashion.

So, here we decide the number of food stalls according to the optimal value of k(number of clusters) and according to that we find the perfect location for the food stalls.

### **Android Studio:**

We used Android Studio for building the app as it is the official Integrated Development Environment (IDE) for Android app development, based on IntelliJ IDEA. It is a unified environment where we can develop for all Android devices. Apply Changes to push code and resource changes to your running app without restarting your app.

## Parts Involved in the Process:

1. Android app with user authentication using firebase Auth, user profile with their location.
2. Firebase Realtime Database to store the locations of the currently located food stalls.
3. Find the nearest food stalls available to the user.
4. Collect the location data from where the requests are coming from.(JSON format data)
5. Converting the JSON data into excel format.
6. Using python to read the data..
7. Using the elbow method to find the optimal number of food stall required(K).
8. Using the K-means algorithm, find the optimal location of the K food stalls.
9. Represent the K food stalls on Google Map.
10. Update the database with new locations of the food stalls.

## SCHEDULE OF THE PROJECT (GANTT CHART)



## **SWOC ANALYSIS**

### **Strengths**

- Effective communication between stakeholders (NGOs) and the people in need.
- Better service than existing methods
- Interactive User Interface
- User Friendly

### **Weaknesses**

- Some people might not have mobile devices.
- People might not know how to use android applications.
- Weak internet signals in certain localities.
- Vulnerability of fake NGOs

### **Opportunities:**

- To do our contribution in helping the poor.
- Scalability - This could be expanded into an International project
- Opportunity to implement Machine learning techniques to solve real life problems.
- Opportunity to use the latest android tools.

### **Challenges:**

- Educating poor people about our application
- Providing technical support
- Recognising fake NGOs, there for malicious reasons
- Implementing machine learning knowledge into android application while development.

## **Discussions / Interactions with Stakeholders**

The stakeholders of this project are :

1. NGOs
2. Food Providers

Impacts and Gains :

1. Reduce the hunger of the people and children.
2. More healthy individuals.
3. Ability to work longer and earn more for adults.
4. Better Development of children on streets in terms of cognitive skills, health etc..

Interactions :

We interacted with orphanages and low wage workers to understand the situation and their problems.

To understand the health and present conditions of children on the street, we talked with orphanages.

To understand about the adults working for low wages and their plight, we talked with low wage workers.

Questions Asked for orphanage organisers :

1. What are the health conditions of the children in the orphanage and how is it different from children on streets?
2. How much food does it take to feed the children in the orphanage? ( To get an idea as to how much we could need to feed children on street)
3. Where do the children on streets present more ? ( what kinds of areas such children exist more)

Questions Asked for Adult low wage workers :

1. Are you from a place within the state or migrated for work?
2. How much do you earn on a day to day basis?

3. How many times a day do you eat food ?
4. What is the type of food you take?
5. Do you receive any kind of support from the government ?
6. How many hours do you work ?
7. What is the type of work you do?
8. How educated are you?
9. Do you have a phone . If so, what kind ?

### Summary of Answers :

Most of the Adults are migrant workers who work as construction workers, plumbing and other low paying jobs. They do not get any support from the government as they are not part of the voting population of the state.

On average they eat 2 meals a day which consists of rice with something . It is not sufficient for them nor for their children and it doesn't give them any essential minerals and vitamins . Most people go to work at 8 am and end work at 6 or 7 pm with breaks for taking food. Almost all the workers didn't even finish 7th grade. Some people have been doing the same kind of work for generations and didn't even go to school. They do not have access to any smart devices nor know how to use them.

The children who are in orphanages are more healthy and are having better access to food than the ones on the street as expected. The people on the street mostly reside in slum areas. They sleep on pavements and bus stops and during daytime they work on traffic lights selling something or begging. Though some NGOs are trying to get them to orphanages and provide them with food it's not enough.

**Interaction with people:**  Tarp Project : Food for neddy

We also Interacted with stakeholders to figure out the challenges they face while supplying services, how they found people in need, ..etc.

**Interaction with stakeholder:**  Tarp Project : Food for neddy

## **LITERATURE SURVEY: -**

Shubham Singh Rana et al. [1] They develop a web-based application that will provide a platform to the people to donate their leftovers along with books and clothes, also the people who are in need of a meal will get something to fill their empty stomach. This platform will be of great advantage to avoid any kind of food wastage and people can also join us as volunteers who can donate food in their neighborhood. They will establish a partnership with NGO's who will be helping us to donate the food, clothes and books to the right people. This system will create a common collaboration portal for hotels/restaurants and NGO's. Hotels/ restaurants, NGO's and users, registered to the admin, all of them will send a request to the admin and admin will fulfill these requests. In this system mainly there are four modules which are admin, NGO's, volunteers and third parties registered to the admin like the user, hotels, restaurants, etc. any individual or an organization trying to donate food through this application will create a new request to the admin and the admin will grant the request whenever it can.

P. Que et al. [2] this paper introduced the hybrid framework and analyzed the pros and cons of native & hybrid mobile app development paradigms. Hybrid apps are easier to develop and maintain, but it comes with the price of performance. Native apps deliver better experience and performance but it cannot be cross platform. Both paradigms are suitable for fulfilling different needs and preferences, so usage of either depends on the individual requirements.

Sandeep Kumar et al. [3] two such major factors are web browser and GPS services. Both of these functionalities are already implemented but are only in the hands of manufacturers not in the hands of users because of proprietary issues, the system does not allow the user to access the mobile hardware directly. But now, after the release of android based open source mobile phone a user can access the hardware directly and design customized native applications to develop Web and GPS enabled services and can program the other hardware components like camera etc. In this paper we will discuss the facilities available in android platform for implementing LBS services (geo-services).

Ahuja et al. [4] this is an introductory chapter to machine learning containing supervised, unsupervised, semi-supervised, and reinforcement algorithms and applications of machine learning. This chapter covered four classification techniques (Logistic Regression, Decision Tree, K-Nearest Neighbors, and Naive Bayes) and K means, and Hierarchical clustering algorithms considering two well-known datasets (Iris and tennis) using Python.

Pham et al. [5] The K-means algorithm is a popular data-clustering algorithm. However, one of its drawbacks is the requirement for the number of clusters, K, to be specified before the algorithm is applied. This paper first reviews existing methods for selecting the number of clusters for the algorithm. Factors that affect this selection are then discussed and a new measure to assist the selection is proposed. The paper concludes with an analysis of the results of using the proposed measure to determine the number of clusters for the K-means algorithm for different data sets.

Trupti m Kodinariya and Prashant Makwana [6] the several clustering algorithm has been proposed. Among them k-means method is a simple and fast clustering technique. They address the problem of cluster number selection by using a k-means approach. They can ask end users to provide a number of clusters in advance, but it is not feasible. End user requires domain knowledge of each data set. There are many methods available to estimate the number of clusters such as statistical indices, variance based method, Information Theoretic, goodness of fit method etc...The paper explores six different approaches to determine the right number of clusters in a dataset.

Fabian Pedregosa et al. [7] this paper focused on bringing machine learning to non-specialists using a general-purpose high-level language. Emphasis is put on ease of use, performance, documentation, and API consistency. It has minimal dependencies and is distributed under the simplified BSD license, encouraging its use in both academic and commercial settings. Source code, binaries. Importantly, the algorithms, implemented in a high-level language, can be used as building blocks for approaches specific to a use case, for example, in medical imagination.

Moroney, L [8] In this journal we got an introduction to the Firebase Realtime Database – and we saw how it can be used as a fully functioning database that gives, they Realtime capabilities to build a very responsive app. They saw how data was structured as a JSON table, and how with a little rethinking we could perform queries just like those in SQL. They also learned about value event listeners that will execute when data changes, so can update the state of application, or their SQL-like cousin – the listener for single value events that can be used to run queries. They then saw how can find fields within a database and update or delete them! This journal touches on what we can do with the Firebase Realtime DB, but there's a whole lot more to discover.

Chunnu Khawas and Pritam Shah [9] The web application has become more and more reliant upon large amount of database and unorganized data such as videos, images, audio, text, files and other arbitrary types. It is difficult for Relational Database Management System (RDBMS) to handle the unstructured data. Firebase is a relatively new technology for handling large amount of unstructured data. It is very fast as compared to RDBMS. This paper focuses on the application of Firebase with Android and aims at familiarizing its concepts, related terminologies, advantages and limitations. The paper also tries to demonstrate some of the features of Firebase by developing an Android app.

Ashoori et al.[10] This study uses clustering method to describe blood donor behaviors. The goal of this clustering is provide the capability to determine blood donor segmentation for increasing the accuracy of manager predictions. We use k-means clustering algorithm for number of cluster 2, 3, 4, 5 and 6 and then they use the Dunn's index for distinguishing the optimal number of cluster. The results of Dunn's index show that optimal number of cluster is 2 which two-step clustering algorithm confirms it. The results of blood donor clustering describe that men have a tendency to donate their blood

Sudhir Singh and Nasib Singh Gill [11] Study of this paper describes the behavior of K-means algorithm. Through this paper we have try to overcome the limitations of K-means algorithm by proposed algorithm. Basically actual K-mean algorithm takes lot of

time when it is applied on a large database. That's why the proposed clustering concept comes into picture to provide quick and efficient clustering technique on large data set. In this paper performance evaluation is done for proposed algorithm using Max Hospital Diabetic Patient Dataset.

S. Ebenazer Roselin et al.[12] they are providing the trustworthy crowd funding through government, which is the funding of a project by a large number of supporters who can each contribute a small amount. Only authenticated recipient and donor can request and donate money here. This system uses the clustering algorithm to filter the data from large scale of datasets and uses k-nearest neighbour algorithm for clustering the similar data from large datasets. These systems help in automatically notifying the donors according to their interest in donation on any particular day, for example, on their birthday, and appreciate the donors to further improve their sequence of donation.

Anand Sutariya and Kiran Amin [13] The original k-means algorithm is widely used for clustering large sets of data. But it does not always guarantee good results, as the accuracy of the final clusters depend on the selection of initial centroids. Moreover, the computational complexity of the original algorithm is very high because it reassigns the data points a number of times during every iteration of the loop. Here they present an improved k-means algorithm which combines a systematic method for finding initial centroids and an efficient way for assigning data points to clusters. This method ensures the entire process of clustering in time without sacrificing accuracy of clusters.

Mayra Z. Rodriguez et al.[14] they performed a systematic comparison of 9 well-known clustering methods available in the R language assuming normally distributed data. In order to account for the many possible variations of data, we considered artificial datasets with several tunable properties (number of classes, separation between classes, etc). In addition, we also evaluated the sensitivity of the clustering methods with regard to their parameters configuration. The results revealed that, when considering the default configurations of the adopted methods, the spectral approach tended to present particularly good performance. We also found that the default configuration of the adopted implementations was not always accurate. In these cases, a simple approach based on random selection of parameters values proved to be a good alternative to improve the performance. All in all, the reported approach provides subsidies guiding the choice of clustering algorithms.

Carolyn Cordery et al.[15] This paper proposes regulation differentiated according to charities' main resource providers. This could reduce cost and increase the regulator's effectiveness through focusing effort. In addition, this differentiation segments charity types according to the theories that explain why these organisations form and operate. We demonstrate the feasibility of such segmentation by use of cluster analysis of data on New Zealand registered charities and show which charities could benefit from differentiated regulation.

GU Ziwen et al. [16] This paper proposes a method for automatically selecting cluster centers based on Chebyshev's inequality. MG-DPC is implemented on the dataset of load-data to realize load classification. The clustering performance is evaluated using five validity indices compared with four typical clustering methods. The experimental results demonstrate that MG-DPC outperforms other comparison methods.

Gauldie et al. [17] They present two case studies of prototype implementations of browser-based visual analytics tools leveraging the Louvain clustering algorithm. These are implemented in ApertureJS, a new open source, JavaScript visual analytics library. Two large datasets—CharityNet and Bitcoin—were used to test aggregation of large networks to reveal communities and answer analytical questions. Implementation performance is of interest to support interactive visualizations for time-sensitive, actionable analysis. The CharityNet application visualizes a root charity, all the donors who have donated to that charity, and all the other charities to which they have donated. Louvain aggregation was performed on these subsets, ranging in size from just a few nodes and links to ones with approximately 200K nodes and 240K links. With this single-threaded implementation of Louvain aggregation running on a single four-core processor with 24 GB of RAM, processing time was recorded for each subset. While processing was not completed for the entire dataset, we estimate it would have required approximately 45 hours of continuous processing time for the entire dataset of 1.8M nodes (1.6 GB).

L. M. Naeni et al. [18] In this study they analyse complete networks derived from field survey and market research through proposing an efficient methodology based on proximity graphs and clustering techniques enhanced with a new community detection algorithm. The specific context is the charity and Not-For-Profit sector in Australia and consumer behaviours within this context. To investigate the performance of this methodology they conduct experiments on the network extracted from a dataset that contains responses of 1,550 individual Australians to 43 questions in a quantitative survey conducted on behalf of the Australian Charities and Not-for-Profits Commission to study the public trust and confidence in Australian charities. Here, we generate the distance matrix by computing the Spearman correlation coefficient as a similarity metric among individuals. Then, several types of k-Nearest Neighbour (kNN) graphs were calculated from the distance matrix and the new community detection algorithm detected groups of consumers by optimizing a quality function called "modularity". Comparison of obtained results with the results of the BGLL algorithm, a heuristic given by the publicly available package Gephi and the MST-kNN algorithm, a graph-based approach to compute clusters that has several applications in bioinformatics and finance, reveals that our methodology is effective in partitioning of complete graphs and detecting communities. The combined results indicate that behavioural models that investigate trust in charities may need to be aware of intrinsic differences among subgroups as revealed by their analysis.

Lin. T. et al. [19] in this work, a novel hierarchical clustering framework is proposed for facility location selection, which can flexibly support a wide variety of optimisation targets and the combinations of multiple practical constraints that are vital in the real logistics scenarios. Beyond the original hierarchical clustering algorithm, it incorporates a looking-forward mechanism that alleviates the 'greedy trap' by utilising global information. These advantages enable the proposed method to generate reliable solutions with high time efficiency. As demonstrated by the experimental results on real JD Logistics data, the proposed method outperforms the widely adopted GGA and VNS algorithms. It also has a much lower computation cost compared to the SCIP solver, while the quality of solutions are within an acceptable range.

Xing Wang et al. [20] this paper innovatively proposes a highly efficient traffic outlier detection framework based on the study of road traffic flow patterns. The main research works are as follows: (1) data pre-processing, the road traffic flow matrix of the roads is

calculated based on the collected GPS data, the non-negative matrix factorisation algorithm is chosen to reduce the dimension of the matrix. (2) Road traffic flow pattern extraction, the fuzzy C-means clustering algorithm with the Optimal k-cluster centre (K-FCM) is adopted to cluster the roads with the same road traffic flow pattern. (3) Outlier detection model training and evaluation, kernel density estimation is introduced to fit the probability density of roads traffic flow matrices which are used to train the back propagation neural network based on particle swarm optimisation to obtain the outlier detection and evaluation model, and a threshold is introduced to optimise the precision and recall of the model. The experimental results show that: the average precision and recall of the proposed method in this paper are 95.38% and 96.23%, respectively, and the average detection time is 28.4 seconds. The method has high accuracy, high efficiency and good practical significance

Andrea Vocino et al. [21] The purpose of this paper is to seek to assess whether online commercial panel volunteering can be segmented based on their motivations, using the volunteer functions inventory. The authors also investigate whether segments exist which differ in demographic characteristics.

Laura R et al. [22] Prior research considers the extent to which public assistance recipients' charitable activity differs from the habits of the general population. Although receiving public assistance is negatively associated with donating money, the relationship to volunteering is unclear. In response to challenges overcoming selection bias, they conducted a multivariate cluster-based subgroup analysis to reduce bias in our claims about the ways in which public assistance receipt affects charitable activity. This innovative approach to dealing with the problem of selection bias has implications and applications across the social sciences.

J. Sylvester et al. [23] in this paper, they investigate the relationship between social media data, web donations, and traditional media coverage. We start by examining the temporal relationships, then move on to spatial relationships and patterning. Finally, they explore spatiotemporal patterns, examining the confluence of these patterns. Their results indicate that social media can be used as an indicator of both spatial and temporal patterns of real world events such as donations to a large not-for-profit. Their models can be used to make recommendations to managers about how to target marketing messaging appropriately given the current state of social media. [24] They demonstrated the global k-means algorithm, an incremental approach to clustering that dynamically adds one cluster centre at a time through a deterministic global search technique consisting of  $N$  (where  $N$  is the size of the data set) k-means algorithm executions from suitable initial positions. They also offer changes to the method that would lower the computational load without compromising the quality of the solution. The proposed clustering approaches are compared favourably to the k-means algorithm with random restarts on well-known data sets. [25] IDBSCAN outperforms DBSCAN in terms of efficiency. For data clustering, IDBSCAN can also identify arbitrary forms and detect noisy points. This study introduces a new technique based on the IDBSCAN concept, in which K-means is used to locate high-density centre points, and IDBSCAN is then used to expand clusters from these high-density centre points. IDBSCAN has a faster execution time since it selects representative locations in seeds to speed up the process. According to the simulation, the proposed KIDBSCAN produces better accurate clustering results. In addition, this innovative strategy lowers

the cost of I/O. DBSCAN and IDBSCAN are outperformed by KIDBSCAN. [26] Gait analysis is now considered a promising contribution to early detection of cognitive and physical decline in the elderly. However, recent attempts at indoor gait analysis methods are limited because they use only average walking speeds. The application of density-based clustering algorithms to indoor location datasets can accelerate context awareness in gait analysis and thereby improve the quality of information on serious gait disorders. This article introduces the application of DBScan, a well-known knowledge detection algorithm, to indoor Kinect location records collected at the Active and Healthy Aging Laboratory at the Laboratory of Medical Physics at the University of Aristotle in Thessaloniki. The aim of this article is to provide evidence that such an approach can effectively distinguish high-density areas in indoor spaces, which can then be contextually transferred to datasets from actual parents' homes based on gait analysis.[27] By analysing the performance of the two policies in a formal model, this research tried to establish the conditions under which monetary aid is likely to be more or less effective than direct food distribution. Before going over the paper's findings, it's important to note that the model assumes that the policy goal is to simply reduce predicted mortality.

As a result, the findings do not represent the differences in the effects of the two strategies on other factors besides mortality. They don't consider, for example, how financial relief might help keep the famine region's economic infrastructure alive by maintaining a demand for trade and transportation.

.The findings of this study imply that the relative usefulness of cash and direct food aid will be highly dependent on trader behaviour and whether food is exported, imported, or neither exported nor imported. If food is exported from the famine area, cash relief is proven to be the best option if traders behave competitively.[28] The typical way to design and develop a mobile application is to sketch the graphical user interface (GUI) for the various screens within the application, and then create the actual GUI from those sketches. This includes identifying which layout to use, which widgets to add, and how to configure and connect the various parts of the GUI. To help with this difficult and time-consuming task, we offer GUIFetch, a technique that takes application sketches as input and leverages the growing number of open source applications in public repositories to identify applications with a similar GUI and transition to the included application Sketch. GUIFetch first searches public repositories to find potential applications using keyword matching. It then creates an on-screen model of the identified application and transitions on-screen, uses a combination of static and dynamic analysis, and calculates indicators of similarity between the model and the provided sketch. Finally, GUIFetch ranks the identified applications (or parts of them) based on a calculated similarity score and creates a visual rating of the results along with the respective application code. We implemented GUIFetch for Android applications and evaluated it through user surveys with various types of applications.[29] IP network engineering, administration, and control, as well as other essential disciplines, benefit greatly from network traffic classification and application identification. Current common methods, such as port-based and payload-based, have flaws, and the machine learning-based method has the potential to overcome them. The

traffic is categorised based on statistical features that are not dependent on the payload. This paper discusses the many layers of network traffic analysis as well as important machine learning information, as well as the issues with port-based and payload-based traffic classification approaches. Taking into account the importance of the machine learning-based method.

To evaluate the efficiency and performance of unsupervised K-means, we conduct experiments. We use feature selection and log transformation to improve accuracy and locate the best feature set. The experimental findings on several datasets show that the approach may achieve up to 80% overall accuracy, with the accuracy improving to 90% or higher following a log transformation. [30] The problem of estimating the exact number of clusters is a big and difficult problem in cluster research. There are different methods

for this in hierarchical clustering; A typical approach is to try to group a number of different clusters and compare them against a metric to estimate the number of clusters. On the other hand, there is no such method to automatically estimate the number of clusters in agglomerative hierarchical clustering (AHC) because AHC generates cluster families with different number of clusters simultaneously in the form of a dendrogram. One exception is Newman's method in network clusters, but this method has no useful dendrogram output. The purpose of this article is to propose a new method to automatically estimate the number of clusters in AHC. To this end, we demonstrate two approaches, on the one hand the variance of the cluster validity measure and on the other hand a method for selecting statistical models such as BIC.

## Table



= Used



= Not Used

Research papers	K means algorithm:	Mean -Shift Clustering:	Density-Based Spatial Clustering	Expectation–Maximization Clustering using Gaussian Mixture Models (GMM):
1.	NO CLUSTER	NO CLUSTER	NO CLUSTER	NO CLUSTER
2 USED HYBRID	✗	✗	✗	✗

3 LBS	✗	✗	✗	✗
4 Hierarchical clustering	✓	✗	✗	✗
5	✓	✗	✗	✗
6	✓	✗	✗	✗
7	✓	✗	✗	✗
8 Firebase Realtime DB	✗	✗	✗	✗
9 Firebase API	✗	✗	✗	✗

10	✓	✓	✓	✓
11	✓	✗	✗	✗
12	✓	✗	✗	✓
13	✓	✗	✗	✗
14	✓	✓	✓	✓
15	✓	✓	✓	✗
16	✓	✗	✓	✓
17 Louvain Clustering	✓	✗	✗	✗
18 MST-kNN algorithm	✓	✗	✗	✓
19 Hierarchical clustering	✗	✗	✗	✗
20	✓	✗	✓	✗
21 ANOVA and $\chi^2$	✗	✗	✗	✗
22	✗	✗	✗	✗
23 Spatiotemporal Cluster Analysis	✗	✗	✗	✗

S.No	Title of the paper	Author	Year	Pros and Cons
1.	A Comprehensive Comparison Between Hybrid and Native App Paradigms.	Peixin Que, Xiao Guo, Maokun Zhu	2016	<b>Pros:</b> rapid development of HTML5 and hybrid app frameworks, the performance of hybrid app could be vastly improved <b>Cons:</b> cross platform dependency is not up to the mark
2.	GET A BITE: The charity-based application through which people can donate food, books and clothes to needy people.	Shubham Singh Rana, Satvik Maheswari, Shiv Kumar, Ms. Jyoti Thakur.	2020	<b>Pros:</b> application will reach to all the people of India, n can play a major role to help India become more developed in coming future by making all the citizen of India happy and prosperous <b>Cons:</b> local people who should access this app have very less knowledge about smart devices
3.	Location based services in Android.	Ch. Radhika Rani, A. Praveen Kumar, D. Adarsh, K. Krishna Mohan, K.V.Kiran	2012	<b>Pros:</b> utilize geographical position of the mobile device, they can utilize multiple technologies such as GPS, WIFI... etc <b>Cons:</b>

				The whole centralized system which is providing services is not optimized as every service request is sent through a common domain
4.	Classification and Clustering Algorithms of Machine Learning with their applications.	Ravinder Ahuja, Aakarsha Chug, Shaurya Gupta, Pratyush Ahuja, Shruti Kohli	2020	<b>Pros:</b> It covered all latest classification techniques. i.e., Logistic Regression, Decision Tree, K-Nearest Neighbors, and Naive Bayes
5.	Selection of K in Kmeans clustering.	D T Pham, SS Dimov, and C D Nguyen	2005	<b>Pros:</b> proposed method can suggest multiple values of K to users for cases when different clustering results could be obtained with various required levels of detail. <b>Cons:</b> computationally more expensive with larger data
6.	Review on Determining of Cluster in K-means Clustering	Trupti m Kodinariya, Prashant Makwana	2013	<b>Cons:</b> clusters are not in data but in the viewing eye, clusters are not completely

				conventional in data
7.	Scikit-learn: Machine Learning in Python.	Olivier Grisel, Mathieu Blondel	2011	<b>Pros:</b> it relies on the scientific Python ecosystem; it can easily be integrated into applications outside the traditional range of statistical data analysis. <b>Cons:</b> not optimal for very large datasets
8.	The Firebase Realtime Database.	Moroney, L	2017	<b>Pros:</b> dynamic and optimised control over the database, easy accessibility <b>Cons:</b> expensive for maintaining large number of instances
9.	Application of Firebase in Android App Development-A Study.	Chunnu Khawas, Pritam Shah.	2018	<b>Pros:</b> dynamic and optimised control over the database, easy accessibility <b>Cons:</b> expensive for maintaining large number of instances
10	Using Clustering Methods for Identifying Blood Donors Behavior.	Ashoori, Maryam & Taheri, Zahra.	2013	<b>Pros:</b> used Dunn's indexing to show optimal clusters <b>Cons:</b> blood donations by

				people are uninterpreted
11	Analysis and Study of K-Means Clustering Algorithm, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 02, Issue 07 (July 2013)	Sudhir Singh And Nasib Singh Gill,	2013	<p><b>Pros:</b> In K-means clustering algorithms, the number of clusters (k) needs to be determined beforehand but in proposed clustering algorithm it is not required. It generates number of clusters automatically.</p> <p><b>Cons:</b> The compactness of the algorithm can be improved using nearest neighbour algorithm</p>
12	Online Donation Based Crowdfunding using Clustering and K-Nearest Neighbor Algorithm. Volume 10 Issue No.3	S. Ebenazer Roselin, Deepa. R, S. Indira, R. Priya, (2020)	2020	<p><b>Pros:</b> The donor retention and donor recurrence is much better in the proposed system than the existing one. Hence enable more funds.</p> <p><b>Cons:</b> Need to improve the existing models for other scenarios, such as traditional charity activities, especially applied to</p>

				survival data with modelling collaborative tasks in some other domains, such as device failure modelling in engineering, predicting student dropout, and prospecting the career development.
13	An Improvement in K-means Clustering Algorithm, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 02, Issue 01	Anand Sutariya, Kiran Amin	2013	<p><b>Pros:</b> The an improved k-means algorithm which combines a systematic method for finding initial centroids and an efficient way for assigning data points to clusters.</p> <p><b>Cons:</b> The time complexity of the improved algorithm can be further reduced.</p>
14	Clustering algorithms: A comparative approach	Rodriguez MZ, Comin CH, Casanova D, Bruno OM, Amancio DR, Costa LdF, et al	2019	<p><b>Pros:</b> The results revealed that, when considering the default configurations of the adopted methods, the spectral approach usually outperformed the other clustering algorithms</p> <p><b>Cons:</b> This suggests that the algorithms are not sensitive to parameter variations for this dataset.</p>

15	DIFFERENTIATED REGULATION: THE CASE OF CHARITIES. WORKING PAPER SERIES Working Paper No. 93 May 2014.	Carolyn Cordery, Dalice Sim, Tony van Zijl	2014	<b>Pros:</b> This could reduce cost and increase the regulator's effectiveness through focusing effort.
16	A Multi-Granularity Density Peak Clustering Algorithm Based on Variational Mode Decomposition. Volume30, Issue4 July 2021, Pages 658-668.	GU Ziwen, LI Peng, LANG Xun, YU Yixuan, SHEN Xin, CAO Min	2021	<b>Pros:</b> The experiment results in an excellent performance of MG-DPC compared to the four clustering methods <b>Cons:</b>
17	Louvain Clustering for Big Data Graph Visual Analytics	Gauldie, David & Langevin, Scott & Schretlen, Peter & Jonker, David & Bozowsky, Neil & Wright, Willia L. M. Naeni, N. J. De Vries, R. Reis, A. S. Arefin, R. Berretta and P. Moscato	2014	<b>Pros:</b> high memory cluster configurations are used for the implementation of the Louvain aggregation algorithm, which significantly improved the performance  <b>Cons:</b> They did not consider batch-processing/real-time thresholds to determine when to

				re-compute by batch.
18	Identifying Communities of Trust and Confidence in the Charity and Not-for-Profit Sector: A Memetic Algorithm Approach	L. M. Naeni, N. J. De Vries, R. Reis, A. S. Arefin, R. Berretta and P. Moscato	2014	<p><b>Pros:</b> proposed algorithm obtains partitions with higher modularity value, more particularly, in dense k-nearest neighbour graphs with k greater than or equal to 4.</p> <p><b>Cons:</b> 1) the only used metric for fitness evaluation is the modularity value.</p>
19	Hierarchical clustering framework for facility location selection with practical constraints. IET Cyber-Phys. Syst., Theory Appl. 1-16 (2021).	Lin. T., et al.:	2021	<p><b>Pros:</b></p> <p>1) The incorporation of novel looking forward mechanism resulted in high time efficiency. 2) Very low computation cost</p> <p><b>Cons:</b></p>
20	A highly efficient framework for outlier detection in urban trafficflow	Wang, X	2021	<p><b>Pros:</b> 1) problem of traffic outlier detection is addressed and a highly efficient</p>

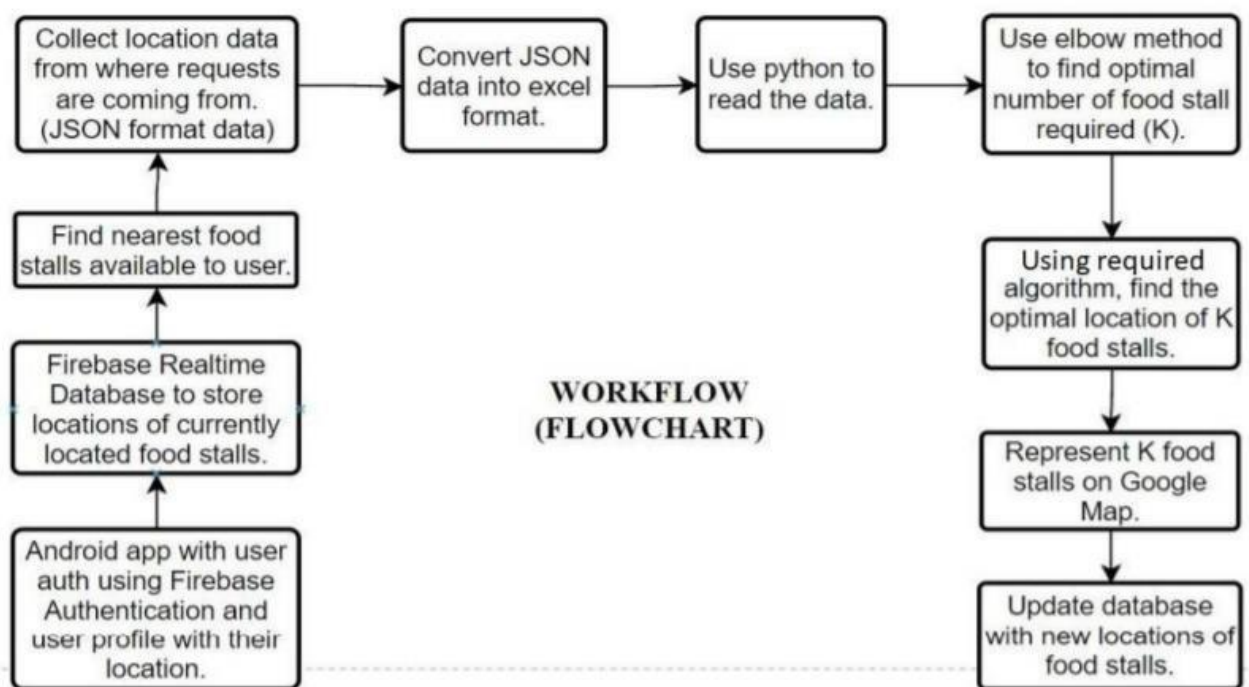
				<p>framework is proposed</p> <p><b>2)</b>compared with the existing method, the proposed method has higher detection accuracy</p> <p><b>Cons:</b> Comparative experiments with pattern-based methods is not done, Doing this would further increase the detection accuracy</p>
21	Segmenting Australian online panellists based on volunteering motivations	Vocino, A., Polonsky, M. and Dolnicar, S	2015	<p><b>Pros:</b> Cluster analysis produced a five-cluster solution, where respondents with low motivations overall comprised the largest grouping</p> <p><b>Cons:</b> The high number of respondents with low motivations may explain the relatively high</p>

				levels of churn that take place within online panels and as a result panel operators would need to continually attract new members
22	A New Strategy for Reducing Selection Bias in Nonexperimental Evaluations, and the Case of How Public Assistance Receipt Affects Charitable Giving	Laura R. Peck, Ida D'Attoma, Furio Camillo, Chao Guo	2012	<p><b>Pros:</b> Mixed classification method is used in clustering to provide better analytical results</p> <p><b>Cons:</b> it does not account for the possibility that those receiving public assistance give of their money and time outside of the system of formal charitable giving</p>
23	Investigating the Spatiotemporal Relationship among Social Media Use, Donations, and Disasters.	J. Sylvester, J. Healey, C. Wang, W. Rand	2021	<p><b>Pros:</b> most blood donation status is uninterrupted that it can be relate to gender of blood donor</p> <p><b>Cons:</b></p>

				1) Very less Optimal number of clusters are obtained
--	--	--	--	--

## METHODOLOGY:

**WORKFLOW** (Parts Involved in the Process): -



**1. Android app with user authentication using Firebase Auth, user profile with their location.**

**2. Firebase Realtime Database to store the locations of the currently located food stalls.**

**3. Find the nearest food stalls available to the user.**

**4. Collect the location data from where the requests are coming from. (JSON format data).**

- 5. Converting the JSON data into excel format.**
- 6. Using python to read the data.**
- 7. Using elbow method to find the optimal number of food stall required (K).**
- 8. Using Mean -Shift Clustering, Expectation–Maximization Clustering, Density-Based Spatial Clustering algorithms, find the optimal location of the K food stalls.**
- 9. Represent the K food stalls on Google Map.**
- 10. Update the database with new locations of the food stall.**

**DATASET: -**

Using the app, we will collect the location coordinates (JSON Format) of the user requesting for the food stall which will be collectively saved in a database and then from the database we will collect the data, convert it into excel format and use it in the algorithm as required

**(Sample Dataset)**

	A	B	C
1	X	Y	
2	-10	-20	
3	10	30	
4	11	28	
5	-12	21	
6	20	-18	
7	-9	-25	
8	12.22	26.98	
9	-9.8	-23.25	
10	10	28.2	
11	9	19.25	
12	-15	-29	
13	9	30	
14	-16	38.4	
15	14	-10	
16	18.11	-25	
17	-10	25.33	
18	-13	40	
19	22	-16	
20	-11.45	30	
21	0	0	
22	0.62	-2.92	
23	-1.11	-2.11	
24	-2.88	3.14	
25	2.45	0	
26	20	45	
27	21.99	48.5	
28	23.33	50.85	
29	19.1	54.22	
30	24.77	55.55	
31	29.99	17.88	
32	-10	65	
33	-11	68	
34	-11.88	64	
35	-9.88	69.8	
36	-8.88	70	
37	-1.33	39.99	
38	-7	63.33	
39	5	-31.2	
40	3.22	-36.22	
41	2.1	-32.11	
42	6.99	-37.33	
43	4.55	-35.1	
44	1.11	-34	
45	18	-9	
46	19	-7	
47	18.66	-16	
48	15.99	-16.11	
49	9.11	70.11	
50	-14.1	32.1	
51	22	16	
52			
53			

## ALGORITHMS USED:

### Mean -Shift Clustering:

Mean shift clustering is a sliding window-based algorithm that attempts to find dense areas of data points. It is a centroid-based algorithm meaning that the goal is to locate the centre points of each group/class, which works by updating candidates for centre points to be the mean of the points within the sliding window. These candidate windows are then filtered in a post-processing stage to eliminate near-duplicates, forming the final set of centre points and their corresponding groups.

### Density-Based Spatial Clustering of Applications with Noise ( ):

The DBSCAN algorithm is based on this intuitive notion of “clusters” and “noise”. The key idea is that for each point of a cluster, the neighbourhood of a given radius has to contain at least a minimum number of points. DBSCAN poses some great advantages over other clustering algorithms. Firstly, it does not require a pre-set number of clusters at all. It also identifies outliers as noises, unlike mean-shift which simply throws them into a cluster even if the data point

is very different. Additionally, it can find arbitrarily sized and arbitrarily shaped clusters quite well.

### **Expectation–Maximization Clustering using Gaussian Mixture Models (GMM):**

With GMMs we assume that the data points are Gaussian distributed; this is a less restrictive assumption than saying they are circular by using the mean. That way, we have two parameters to describe the shape of the clusters: the mean and the standard deviation. This means that the clusters can take any kind of elliptical shape (since we have standard deviation in both the x and y directions). Thus, each Gaussian distribution is assigned to a single cluster.

### **K means algorithm:**

The algorithm takes the unlabelled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. Its Centroid based algorithm, where each cluster is associated with a centroid. Here, the dataset contains the x-coordinate and y-coordinate of the location from where the food stall requests have come and predicts the output which will be considered as the optimal location for the food stall.

### **Elbow method:**

A fundamental step for any unsupervised algorithm is to determine the optimal number of clusters into which the data may be clustered. The Elbow Method is one of the most popular methods to determine this optimal value of k. To determine the optimal number of clusters, we have to select the value of k at the “elbow” that is the point after which the distortion/inertia start decreasing in a linear fashion. So, here we decide number of food stalls according to the optimal value of k(number of clusters) and according to that we find the perfect location for the food stalls.

### **TECHNOLOGIES USED:**

#### **Flutter:**

We used Flutter for building the app as it is the official Google’s framework for app development. It is a unified environment where we can develop for all Android devices. Apply changes to push code and resource changes to your running app without restarting your app.

#### **Google Maps API:**

Google Maps API is a set of application programming interfaces that allows the user to talk to its services. It is one of these smart bits of Google generation technologies that enables us to use take Google Maps take its functionalities and use it directly on our application according to our use. It helps us to upload applicable content material this is beneficial in for our visitors and customize the appearance and experience of the map to fit with the style of our application. In our application we are going to use this API to provide the accurate locations of the food stalls where the food can be deposited by people and provided to the people in need to make it easier for them.

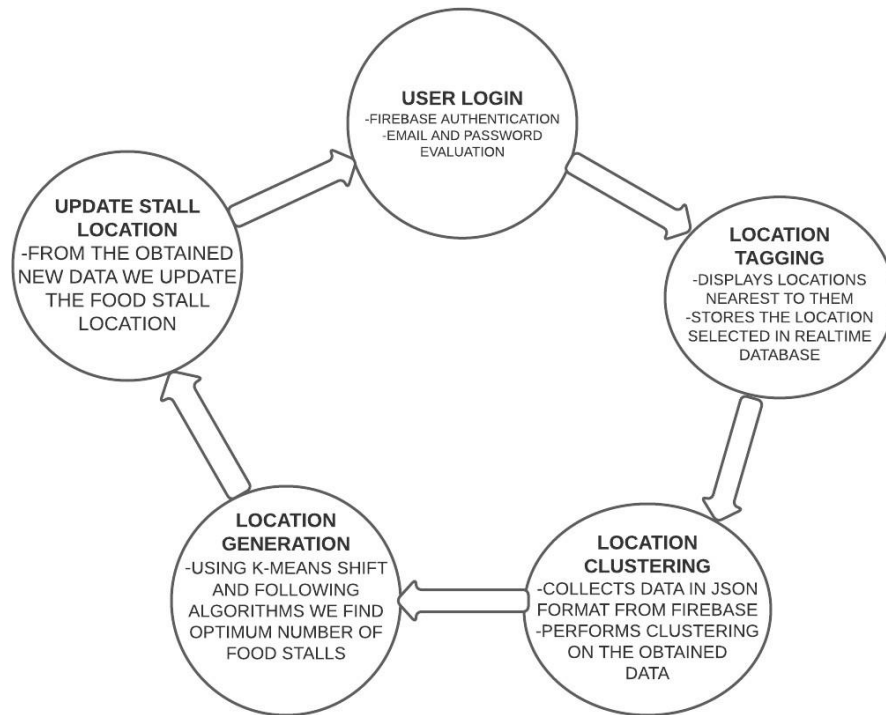
### **Firestore Realtime Database:**

The Firestore Realtime Database is a NoSQL database from which we are able to store and sync the information among our customers in real-time. It is a massive JSON object which the developers can control in real-time. So, in our application we are going to use real time database since the requests of food and new food stalls to det up will keep coming and the application can be refreshed within regular intervals of time so as to find the new store locations or change the previous location according to customer requirements and help the maximum people possible.

### **Firestore Authentication:**

Firestore Authentication helps us to make a secure, stable authentication system easy, while enhancing the sign-in and onboarding experience for end users. It offers an end to-end identification solution, supporting email and password accounts, phone authentication, and applications like Google, Twitter, Facebook, and more. It presents backend services, easy-to-use SDKs, and ready-made UI libraries to authenticate customers on our app. It helps authentication the use of passwords, phone numbers, famous federated identification providers like Google, Facebook and Twitter, and more.

### **ARCHITECTURE:**



## Functional Analysis:

**User Login:** User logs in to the app. The app evaluates the user information with the database and logs the user into his account. If the user information is faulty, then the app immediately displays the message to try again.

**Location Tagging:** Users select location near to them. Now the app stores the information in firebase database and we can extract the json files from firebase.

**Location Clustering:** We collect the data in json format from firebase and convert it into excel format. This excel sheet is the input for our K Means clustering algorithm

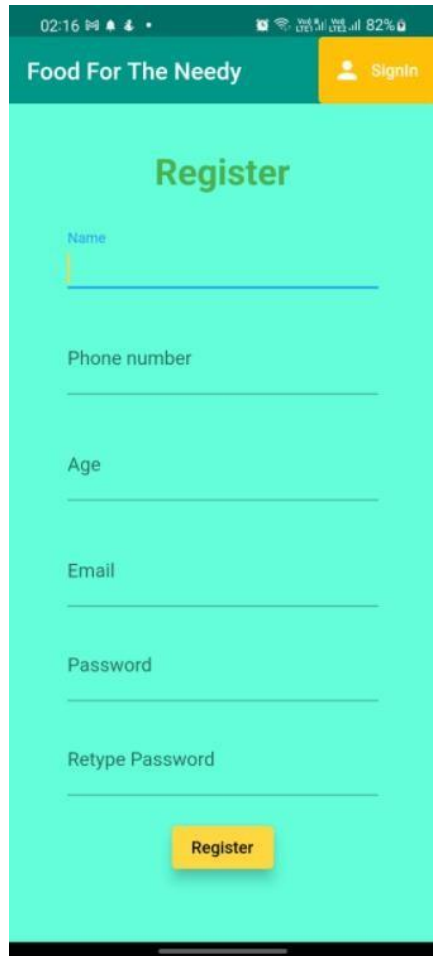
**Location Generation:** The Data is inputted into the clustering algorithm and the clusters are formed. We chose 3 prior cluster counts to be made for efficient location generation. Now using Mean Shift Algorithm, we will obtain the optimal number of food stalls per location and accommodate accordingly.

**Update Stall Locations:** Using new food requests we can forge an updated data set of locations for clustering purposes in frequent interval of times. This new dataset will be used for the next meal.

## **Android App**

### **1) Register**

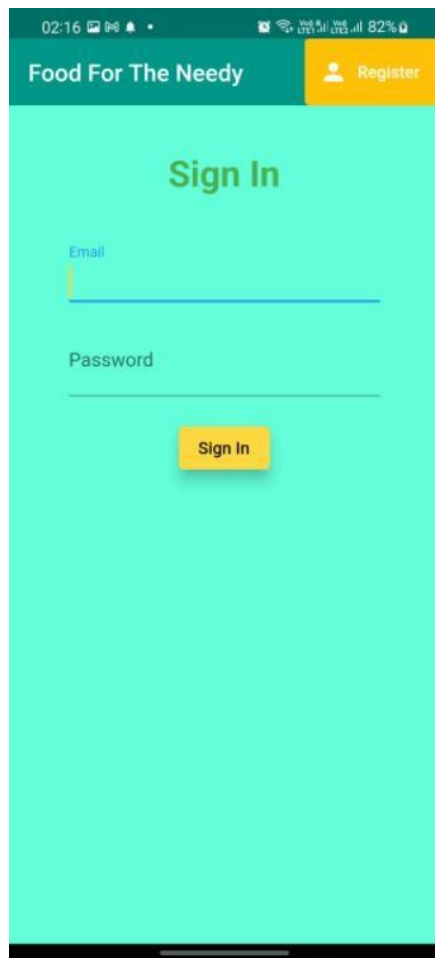
The app starts with the register page if the user is not signed in. In this screen the app is provided with the user details which will be stored in the firebase database.



The screenshot shows the 'Register' screen of the 'Food For The Needy' app. The app's header is teal with the title 'Food For The Needy' and a yellow 'Signin' button with a user icon. The main background is light blue. The 'Register' title is in bold green. Below it are input fields for 'Name', 'Phone number', 'Age', 'Email', 'Password', and 'Retype Password', each with a blue underline. A yellow 'Register' button is at the bottom. The status bar at the top shows the time 02:16, battery at 82%, and various connectivity icons.

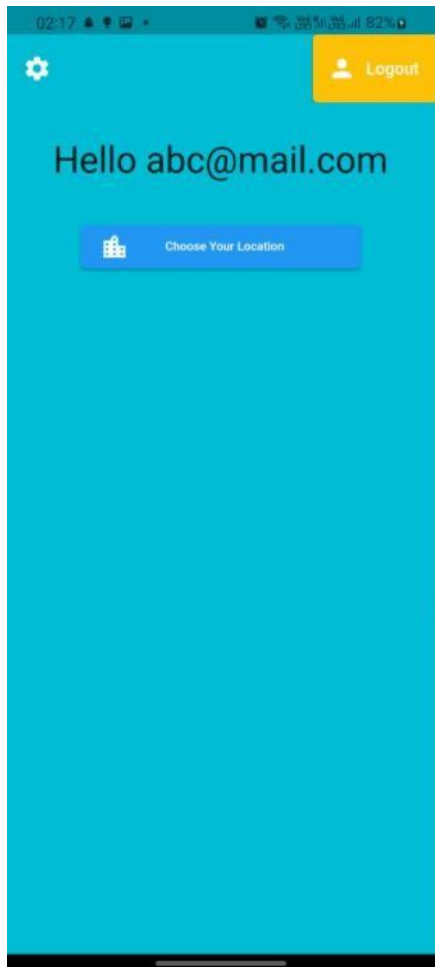
### **2) Sign in**

If the user already has an account, he can sign in to the app using this screen in the app.

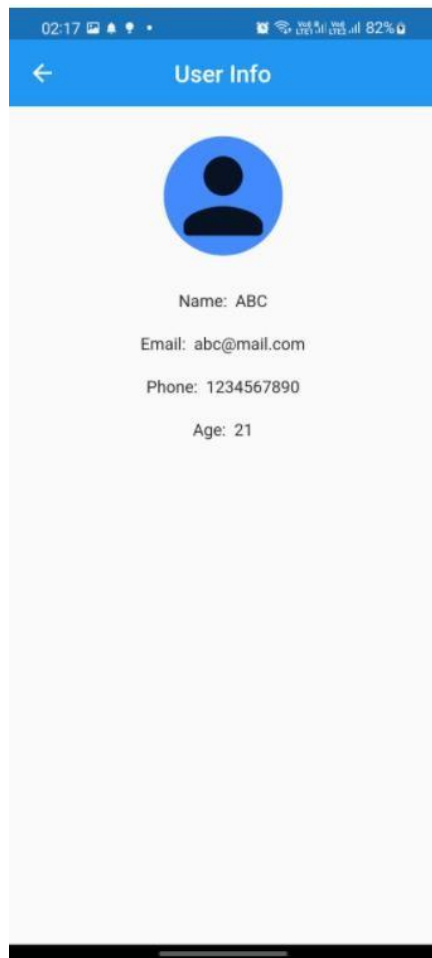


### 3) Home Page

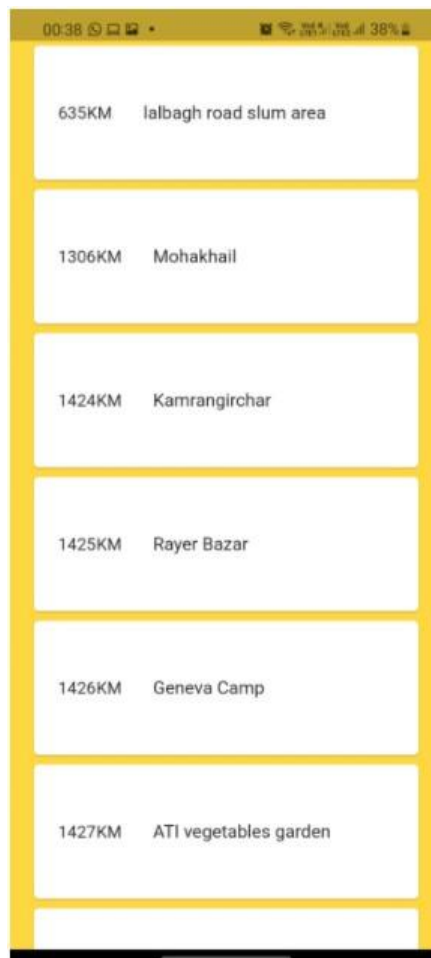
This is the home screen, in this screen the user can access User Info screen, he can choose location of food stall by clicking on choose your location button and he can sign out of the app using logout button.



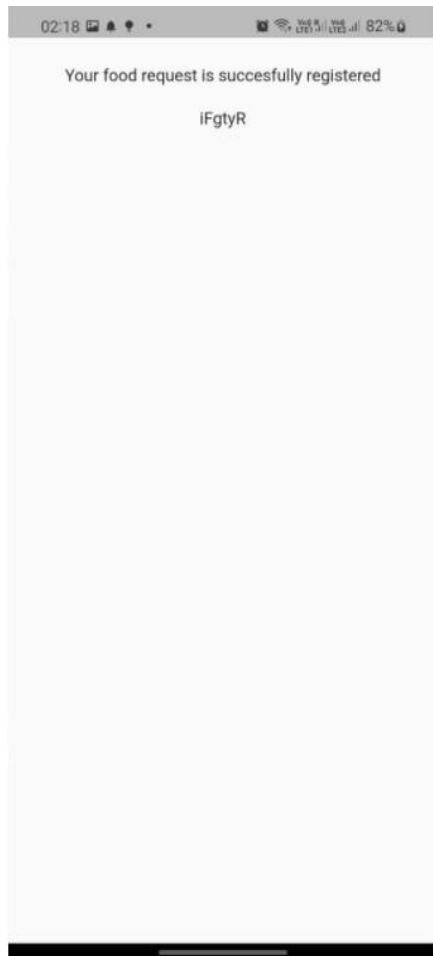
Selecting the bolt symbol the user will be taken to the userdetails screen where he can check his details



This screen will show up once the user clicks on the choose your location button, from this screen he can choose the location of the food stall. Geolocator dependency used in this app made it possible find the distance between the user's current location and the location of the food stall. (Distances are far away since the user is in Hyderabad and locations are in Bangladesh)



After selecting a location the user will be redirected to the food token page where the user can check his food token number.



Firestore Console:

## Location Details stored in database :

<div> <div> <div>🏠</div> <div>&gt;</div> <div>LocationDB</div> <div>&gt;</div> <div>3bP1ln2itiT0mncr97EpT2AGT2r1</div> </div> </div>		
<div> <div>📶</div> <div>foodver1</div> </div>	<div> <div>📁</div> <div>LocationDB</div> </div>	<div> <div>📄</div> <div>3bP1ln2itiT0mncr97EpT2AGT2r1</div> </div>
<div> <div>+</div> <div>Start collection</div> </div>	<div> <div>+</div> <div>Add document</div> </div>	<div> <div>+</div> <div>Start collection</div> </div>
<div> <div>LocationDB</div> <div>&gt;</div> </div>	<div> <div>3bP1ln2itiT0mncr97EpT2AGT2r1</div> <div>&gt;</div> </div>	<div> <div>+</div> <div>Add field</div> </div>
<div> <div>Users</div> <div>stallsDB</div> <div>userCodeDB</div> </div>	<div> <div>K2behogKG0aoQhiNqr3VQRRM2y52</div> <div>LKtUASSa2hPm9AJadtuWFBkj0Sx2</div> <div>MNrsJCirJKUZaN4epSyzdPpYmj33</div> <div>RAdHTi15YAMyJxI1PY207NZJs9f1</div> <div>tU1C3ck6zWPz7BzS7xAZStCEGrj1</div> <div>yR2C3PanxQb87nFGaPQWc8U0JdP2</div> </div>	<div> <div>chosen_lat: 29.951391</div> <div>chosen_lon: 75.813315</div> <div>curr_lat: "28.4277939"</div> <div>curr_long: "77.3227467"</div> <div>locationName: "Punjab"</div> <div>uid: "3bP1ln2itiT0mncr97EpT2AGT2r1"</div> </div>

## User Details Stored in Database :

Users > 0ZamGXIBQ3Z...		
foodver1	Users	0ZamGXIBQ3ZWcw6FW8XmhWNGSdy2
+ Start collection	+ Add document	+ Start collection
LocationDB	0ZamGXIBQ3ZWcw6FW8XmhWNGSdy2	+ Add field
Users	3bP1ln2itiT0mncr97EpT2AGT2r1	age: 21
stallsDB	K2behogKG0aoQh1Nqr3VQRRM2y52	email: "satya@gmail.com"
userCodeDB	LKtUASSa2hPm9AJadtuWFBkj0Sx2	name: "Satya"
	MNrsJCiRjKUZaN4epSyzdPpYmj33	phone: "6309613121"
	Qs0y0aq2WKNf3E18PW69jgxnrm2	
	RAdHTi15YAMyJxI1PY207NZJs9f1	
	f0TZjGTyWdUR8HgmNX0xIFpmv4x2	
	ksscZerp34eLoWsd1jwFGYw7Xk12	
	tU1C3ck6zWPz7BzS7xAZStCEGrj1	
	yR2C3PanxQbB7nFGaPQWc8U0JdP2	

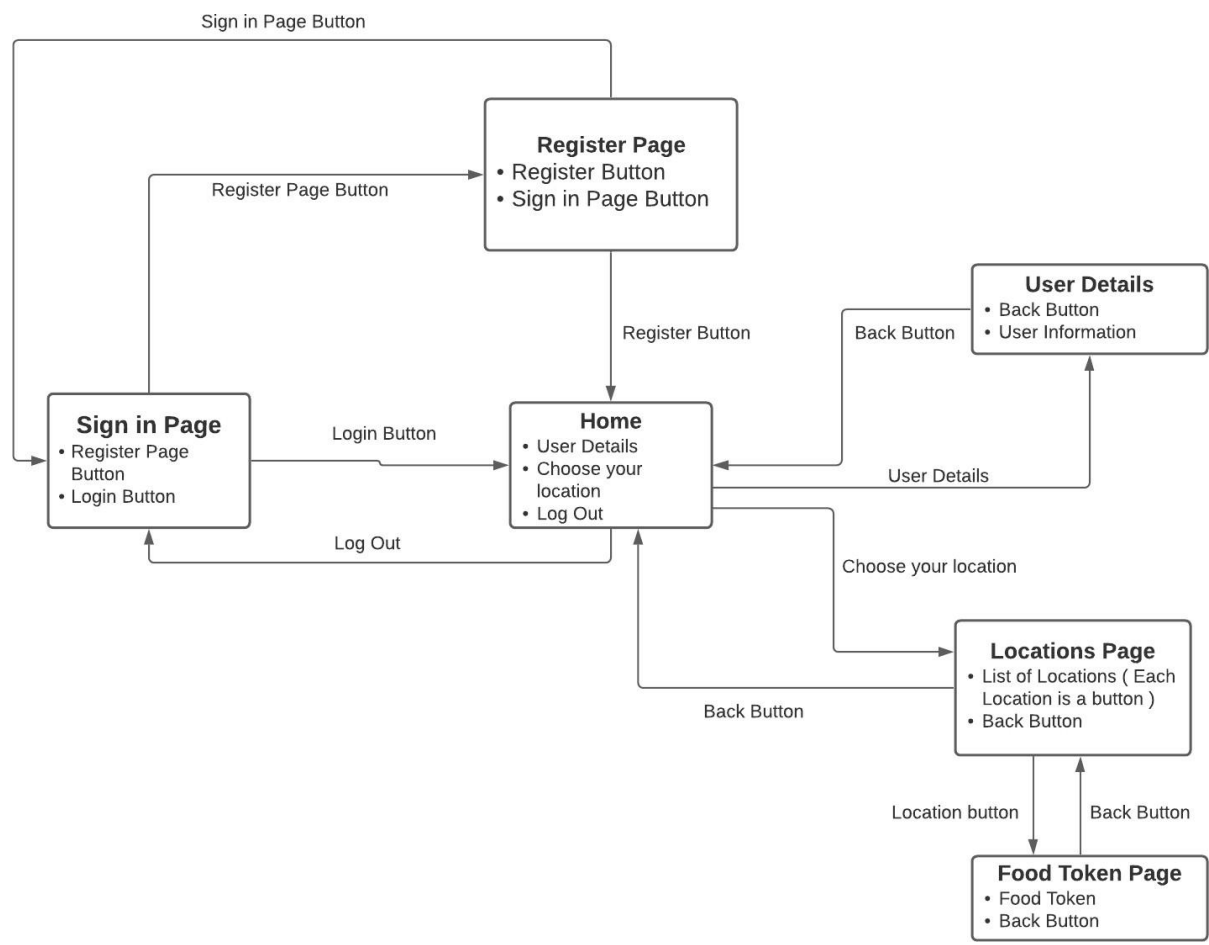
## Food Stall Details Stored in Database :

stallsDB > 1		
foodver1	stallsDB	1
+ Start collection	+ Add document	+ Start collection
LocationDB	1	+ Add field
Users	10	LocationName: "ATI vegetables garden"
stallsDB	11	address: "address 1"
userCodeDB	12	lat: 23.77139
	13	lon: 90.370589
	14	
	15	
	16	
	17	
	18	
	19	
	2	
	20	
	3	

## Food Token Details Stored in Database :

userCodeDB > 3bP1ln2itiT0mn...		
foodver1	userCodeDB	3bP1ln2itiT0mncr97EpT2AGT2r1
+ Start collection	+ Add document	+ Start collection
LocationDB	3bP1ln2itiT0mncr97EpT2AGT2r1	+ Add field
Users	K2behogKG0aoQh1Nqr3VQRRM2y52	foodToken: "W4hKN8"
stallsDB	LKtUASSa2hPm9AJadtuWFBkj0Sx2	uid: "3bP1ln2itiT0mncr97EpT2AGT2r1"
userCodeDB	MNrsJCiRjKUZaN4epSyzdPpYmj33	
	RAdHTi15YAMyJxI1PY207NZJs9f1	
	tU1C3ck6zWPz7BzS7xAZStCEGrj1	
	yR2C3PanxQbB7nFGaPQWc8U0JdP2	

**FLOWCHART FOR THE USAGE OF APP :**



**NGO side of the application :**

**Dashboard :**



**(On Clicking Get Food Tokens Data)**

**All the tokens generated will be visible in an excel sheet**

The screenshot shows a mobile application interface with a dark theme. At the top, the status bar displays the time 00:12, battery level at 40%, and various connectivity icons. Below the status bar, the app title "FoodTokens" is visible in green. A navigation bar contains icons for back, edit, search, save, undo, share, and a menu. The main area displays an Excel spreadsheet with a formula bar at the top showing "fx" and "Name". The spreadsheet has columns labeled A through F and rows numbered 1 through 35. The data is as follows:

	A	B	C	D	E	F
1	Name	User Email	Token number			
2	aaditri	aaditritrimit	W4hKN8			
3	arpit	arpit@gmail	j20lrG			
4	anamika	anamikag	XDvfi0			
5	ABC	abc@mail	IFgtyR			
6	Sandeep	sandeep@	WwYY2G			
7	Md Ibnul f	mdibnulka	JKpnok			
8	Satya	Satya@gh	JMtfPV			
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						

At the bottom of the screen, there is a tab labeled "Sheet1" and a plus icon for adding new sheets.

**(On Clicking Get Food Locations Data)**

**All the order locations generated will be visible in an excel sheet**

The screenshot shows a mobile application interface with a dark theme. At the top, the status bar displays the time 00:12, signal strength, and 40% battery. The app title 'FoodRequests' is centered at the top. Below the title is a navigation bar with icons for back, edit, search, save, undo, share, and a menu. The main content area displays a spreadsheet with the following data:

	A	B	C	D	E	F
1	Name	Lat	Lon	User Email		
2	aaditri	19.951391	75.813315	aaditrimittal123@gmail.com		
3	arpit	17.35846	38.354455	arpit@gmail.com		
4	anamika	17.431271	78.522932	anamikaguha33@gmail.com		
5	ABC	17.390024	78.478174	abc@mail.com		
6	Sandeep	22.705515	75.846945	sandeep@mail.com		
7	Md Ibnul I	20.597825	81.903805	mdibnulkaf1065@gmail.com		
8	Satya	17.390024	78.478174	Satya@gmail.com		
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						

At the bottom, there is a tab labeled 'Sheet1' and a '+' icon to add a new sheet.

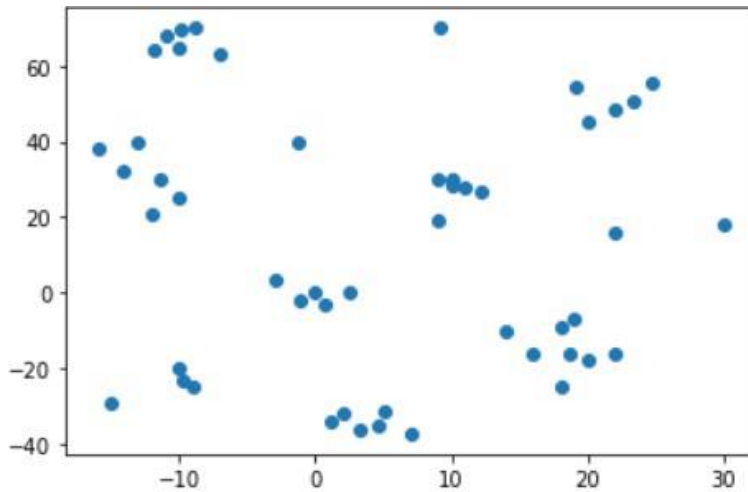
## Experimental Results:

### Experimental Setup:

We used Flutter to develop this app which is a framework provided by Google. We are using the firebase database to store the user data and access it for our machine learning algorithm. We chose firebase because it is a serverless database and more reliable. We used Visual Studio code and Jupyter Notebook for the whole project. We used VS code for app development and Jupyter notebook for machine Learning part.

### Experimental Outcomes:

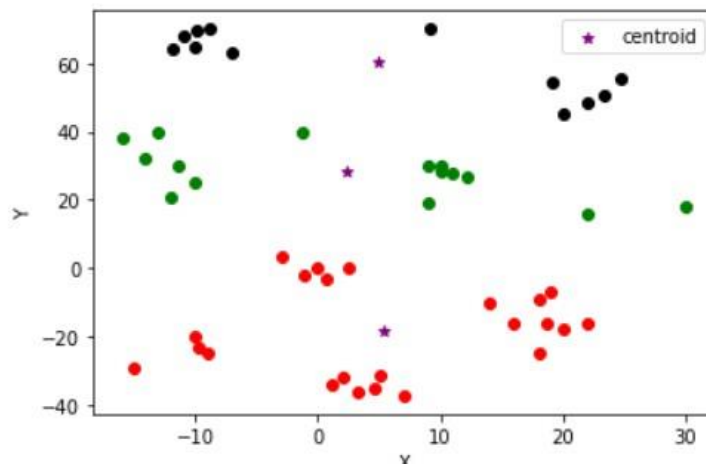
#### Location dataset:



This plot represents the dataset acquired from Firebase Database.

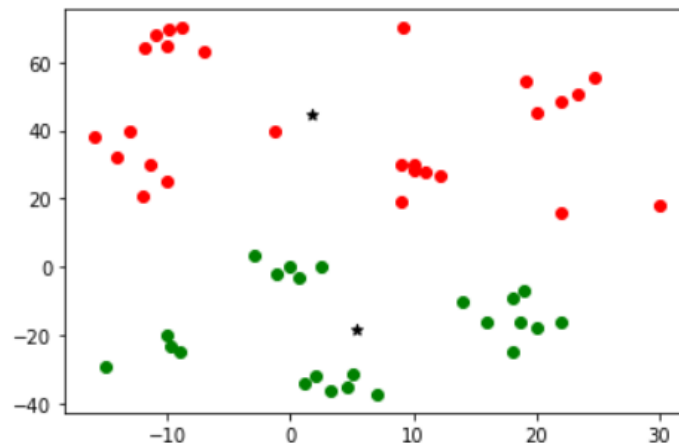
### Clusters formed from the dataset:

Out[17]: <matplotlib.legend.Legend at 0x221065782e0>



The above clusters are formed by applying k means clustering algorithm on the given dataset.

```
[[ 1.72      44.5444    0.96    ]
 [ 5.39173913 -18.35695652  1.    ]]
Estimated clusters: 2
```

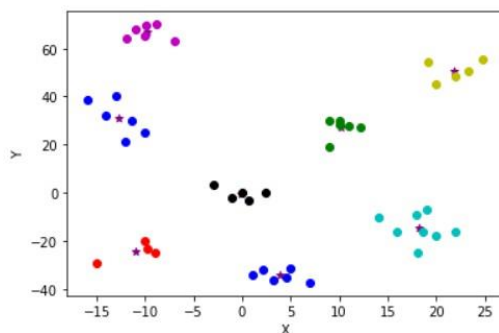


The above optimal clusters are formed by applying mean shift algorithm on the result of k means algorithm. We store this optimal locations data to update the food stall's database frequently.

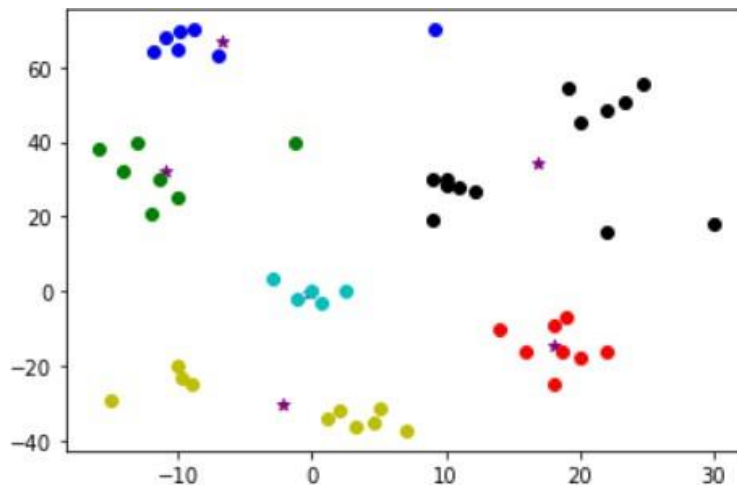
So finally, these optimal clusters will be considered as the locations for food stalls to provide food. These locations are updated in the app for NGOs.

### Clusters obtained by applying DB SCAN algorithm on the same data :

```
Estimated clusters: 8
[-10.95, 10.203333333333333, -12.758333333333333, 18.220000000000002, -0.184, 21.837999999999997, -9.773333333333335, 3.8283333333333336]
[-24.3125, 27.071666666666667, 31.138333333333332, -14.63875, -0.37799999999999984, 50.824, 66.68833333333333, -34.32666666666667]
```



### Clusters obtained by applying GMM algorithm on the same data :



We maintain all this data and make decisions according to the results to change the locations of the food stalls frequently and we think that this is an efficient and user-friendly way to know the optimal locations for the food stalls.

**code:** [Satya5122/food\\_for\\_the\\_needy \(github.com\)](https://github.com/Satya5122/food_for_the_needy)

## **INTERACTION WITH STAKEHOLDERS REGARDING THE PROJECT :**

### **Interaction with the stake holders**

#### **Conclusion:**

We have developed an app which establishes a connection between people in need of food and the NGO's willing to distribute the food to the people in need. The firebase database, being a serverless database is the biggest advantage as this will make the app run and store all the food requests in the database even there is no machine running for the servers. We are also constantly optimising the locations of the food stalls using K means, Mean shift, DB Scan, GMM algorithms. These results provide us with optimised locations data which we can use and change the locations of food stalls in the app frequently or on timely basis. Overall, this fusion of ML algorithms made the task of finding the nearest locations of the food stalls for the people in need easy and efficient.

#### **Future Scope:**

We faced a problem with displaying a route map to the location of food stalls as flutter official plugins are conflicting with each other. The main problem was

Google maps API required a minimum SDK version of 20, but all the other API used in this app requires a minimum SDK version of 21. So, in the upcoming future, this app can be updated so that the user can view the route to the food stall location.

## **References:**

- 1) Que, P., Guo, X., & Zhu, M. (2016, December). A comprehensive comparison between hybrid and native app paradigms. In 2016 8th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 611-614). IEEE.**
- 2) Rana, S. S., Kumar, S., & Thakur, M. J. GET A BITE: The charity-based application through which people can donate food, books and clothes to needy people.**
- 3) Rani, C. R., Kumar, A. P., Adarsh, D., Mohan, K. K., & Kiran, K. V. (2012). Location based services in android. International Journal of Advances in Engineering & Technology, 3(1), 209.**
- 4) Ahuja, R., Chug, A., Gupta, S., Ahuja, P., & Kohli, S. (2020). Classification and clustering algorithms of machine learning with their applications. In Nature-Inspired Computation in Data Mining and Machine Learning (pp. 225-248). Springer, Cham.**
- 5) Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K-means clustering. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 219(1), 103-119.**
- 6) Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. International Journal, 1(6), 90-95.**
- 7) Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.**
- 8) Moroney, L. (2017). The firebase realtime database. In The Definitive Guide to Firebase (pp. 51-71). Apress, Berkeley, CA.**
- 9) Khawas, C., & Shah, P. (2018). Application of firebase in android app**

development-

a study. International Journal of Computer Applications, 179(46), 49-53.

**10) Using clustering methods for identifying blood donors behavior**

**M Ashoori**, Z Taheri - 5th Iranian Conference on Electrical and ..., 2013 - researchgate.net

11. Sudhir Singh And Nasib Singh Gill, 2013, Analysis and Study of K-Means Clustering Algorithm, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 02, Issue 07 (July 2013),

12. S. Ebenazer Roselin, Deepa. R, S. Indira, R. Priya,(2020). Online Donation Based Crowdfunding using Clustering and K-Nearest Neighbor Algorithm. Volume 10 Issue No.3

13. Anand Sutariya, Kiran Amin, 2013, An Improvement in K-means Clustering Algorithm, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 02, Issue 01 (January 2013).

14. Rodriguez MZ, Comin CH, Casanova D, Bruno OM, Amancio DR, Costa LdF, et al. (2019) Clustering algorithms: A comparative approach. PLoS ONE 14(1): e0210236.

15. Carolyn Cordery, Dalice Sim, Tony van Zijl.( 2014). DIFFERENTIATED REGULATION:

THE CASE OF CHARITIES. WORKING PAPER SERIES Working Paper No. 93 May 2014.

16. GU Ziwen, LI Peng, LANG Xun, YU Yixuan, SHEN Xin, CAO Min.(2021). A Multi-Granularity Density Peak Clustering Algorithm Based on Variational Mode Decomposition. Volume30, Issue4 July 2021 ,Pages 658-668.

17. Gauldie, David & Langevin, Scott & Schretlen, Peter & Jonker, David & Bozowsky, Neil & Wright, Willia L. M. Naeni, N. J. De Vries, R. Reis, A. S. Arefin, R. Berretta and P. Moscato, "Identifying Communities of Trust and Confidence in

**the Charity and Not-for-Profit Sector: A Memetic Algorithm Approach," 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, 2014, pp. 500-507, doi: 10.1109/BDCloud.2014.83.m. (2013). Louvain Clustering for Big Data Graph Visual Analytics.**

**18. L. M. Naeni, N. J. De Vries, R. Reis, A. S. Arefin, R. Berretta and P. Moscato, "Identifying Communities of Trust and Confidence in the Charity and Not-for-Profit Sector: A Memetic Algorithm Approach," 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, 2014, pp. 500-507, doi: 10.1109/BDCloud.2014.83.**

**19. Lin. T., et al.: Hierarchical clustering framework for facility location selection with practical constraints. IET Cyber-Phys. Syst., Theory Appl. 1-16 (2021).**

**20. Wang, X., et al. A highly efficient framework for outlier detection in urban trafficflow. IET Intell. Transp. Syst. 2021;1-14.**

**21. Vocino, A., Polonsky, M. and Dolnicar, S. (2015), "Segmenting Australian online panellists based on volunteering motivations", Asia Pacific Journal of Marketing and Logistics, Vol. 27 No. 1, pp. 4-22.**

**22. Laura R. Peck, Ida D'Attoma, Furio Camillo, Chao Guo. (2012). A New Strategy for Reducing Selection Bias in Nonexperimental Evaluations, and the Case of How Public Assistance Receipt Affects Charitable Giving. Volume40, Issue 4 November 2012, Pages 601-625.**

**23. J. Sylvester, J. Healey, C. Wang, W. Rand. (2021 ). Space, Time, and Hurricanes: Investigating the Spatiotemporal Relationship among Social Media Use, Donations, and Disasters. volume 27, p. 467 – 480.**

**24. The global k-means clustering algorithm  
A Likas, N Vlassis, JJ Verbeek - Pattern recognition, 2003 – Elsevier**

**25. KIDBSCAN: A new efficient data clustering algorithm**  
CF Tsai, CW Liu - International Conference on Artificial Intelligence and ..., 2006  
- Springer

**26. Density based clustering on indoor kinect location tracking: A new way to exploit active and healthy aging living lab datasets**  
El Konstantinidis, PD Bamidis - 2015 IEEE 15th International ..., 2015 -  
[ieeexplore.ieee.org](http://ieeexplore.ieee.org)

**27. CASH VERSUS DIRECT FOOD RELIEF**  
Stephen COATE

**28. GUIfetch: supporting app design and development through GUI search**  
F Behrang, SP Reiss, A Orso - ... of the 5th International Conference on ..., 2018 -  
[dl.acm.org](http://dl.acm.org)

**29. Network Traffic Classification using k-means clustering**  
Y Liu, W Li, Y Li - Second international multi-symposiums on ..., 2007 -  
[ieeexplore.ieee.org](http://ieeexplore.ieee.org)

**30. Hierarchical clustering algorithms with automatic estimation of the number of clusters**  
R Abe, S Miyamoto, Y Endo... - 2017 Joint 17th World ..., 2017 - [ieeexplore.ieee.org](http://ieeexplore.ieee.org)