**Data Mining:**

Data mining concept in this application is basically to increase the no of users by doing detailed analysis on the statistics acquired by the application. The statistics will be mapped into data points and data will be converted in to knowledge which will be used in order to increase the sales of the company. The algorithms will track the data points and according to the results new reward system and subscription fess will be decided. The data will be collected by taking into consideration of following points.

1. How much time does a user spends on application?
2. How much time does the user takes to confirm an order?
3. What is the age of the user?
4. How often a user uploads a picture?

By forming data points from this raw data will help us improve our sales by changing it as per the trends and requirements obtained by the mining algorithm.

The simplest form of algorithm we will use is K-Means which gives us the different clusters on the basis of unsupervised learning.

K-Means:

Clustering is the process of partitioning a group of data points into a small number of clusters. For instance, the items in a supermarket are clustered in categories (butter, cheese and milk are grouped in dairy products). Of course this is a qualitative kind of partitioning. A quantitative approach would be to measure certain features of the products, say percentage of milk and others, and products with high percentage of milk would be grouped together. In general, we have n data points xi,i=1...nthat have to be partitioned in k clusters. The goal is to assign a cluster to each data point. K-means is a clustering method that aims to find the positions μi,i=1...k of the clusters that minimize the distance from the data points to the cluster. K-means clustering solves

Arg minc∑i=1k∑x∈cid(x,μi)=arg minc∑i=1k∑x∈ci∥x−μi∥22(formula not copied properly)

where ci is the set of points that belong to cluster i. The K-means clustering uses the square of the Euclidean distance d(x,μi)=∥x−μi∥22. This problem is not trivial (in fact it is NP-hard), so the K-means algorithm only hopes to find the global minimum, possibly getting stuck in a different solution.

References: http://www.onmyphd.com/?p=k-means.clustering.