Identify the Proportion of Benefitted Users from Eazy_Buzy Life App Based on Survey

Neetu Anand*, Dhruvi Goyal**

Abstract: Lives of working women are multifaceted. They need to be assisted for several tasks in their day to day activities. Keeping this in mind, an android based mobile app was created with various features viz. health management, time and stress management, work management, clues about nearest shopping centers, booking cab, food recipes, finding nearest doctor and many more functionalities. Based on the app, the survey form on various key parameters to identify the key users of the app was created and filled by working women, which has been preprocessed first and then after applying clustering, the resultant data is examined and studied.

1. INTRODUCTION

One of the major issues of concern for working women is balancing family related issue with their office work. No matter what the designation and position of women in the office or industry, Indian women have to look after their family along with official work. [1,2] They are treated as a family manager who is going to manage all the work of home like cooking, cleaning and other associated matters of her home and family. For every successful woman there arise needs to prioritize the issues in different sectors viz. professional and personal.

The following are the objective of the research work:

a) To identify the challenging areas of working woman.

b) Building the app for the working women with all the utilities required.

c) To create the survey form based on the app to identify the percentage of benefited users of app.

d) Identify the percentage of working women who are facing glitches in managing time in their personal as well as professional lives.

e) To preprocess the collected data and analyze the result obtained from data clustering.

2. INNOVATION ASPECT IN THE APP

a) Smartphones can't be imagined without mobile apps[3] installed on it, as mobile apps are latest buzz in the field of technology. As per the latest statistics, over 80% of market is owned by android apps, as android is an open source and hence, more innovative, attractive and interactive apps can be built for android platform, or in short, it is an innovative step towards mobile technology.

b) In our app, we ensured innovation by identifying the target users (working women) and devices that would support the app.

c) Our app designed is greatly interactive, which is a step ahead in terms of innovation. With a colorful design and various interactive tools placed within the app, it makes life of working women quite easy.

3. INVENTIVE FEATURES EMBEDDED IN OUR APP

- The app can act as a bodyguard for working women to feel secure and safe.
- Popular recipes are another feature that saves the time of the working women.
- Prioritized task, making shopping easy by locating different stores.
- Providing assistance in health and stress issues, as well as connecting them live with doctors and psychiatrist.

4. ADVANTAGE OF APP

Our system includes all the characteristics to rectify the hindrances faced by working women in current scenario. It aims to reduce the time and money wastage by providing an integrated platform for all the modules [4,5]. The various advantages of the current system are stated as follows:

- Overcomes various glitches faced by working women in their daily's hectic lives.
- Helps in managing lots of work.

**Maharaja Surajmal Institute, GGSIPU; neetuanand@msi-ggsip.org*
***Department of Operational Research, University of Delhi; dhruvigoyal98@gmail.com*

- Interactive and easy to use GUI.
- Access to large number of features in just one or two clicks.
- Efficient utilization of very little available time.
- Easy to operate and maintain.

5. DATA COLLECTION AND ANALYSIS

a) Data Collection and Preprocessing

A survey was conducted amongst the Working Women of different age groups belonging to various states of India like Delhi, Karnataka, Chennai, Tamil Nadu, Uttar Pradesh, Kerala, etc. It consisted of total 22 questions. A total of 269 responses were gathered using the Google Forms. The collected data was stored in the spreadsheet in ".csv" format, so as to perform data mining and analysis on it.

Data preprocessing involves giving one-word variable name to each question so as to reference each field with its specific name in data mining process [6]. It also includes converting all the nominal values of the attributes to numeric values (between 0 and 1), so as to ease the task of data mining process.

From the gathered data, the relevant fields were selected by the feature selection process and gathered in 2 different tables, so as to perform data mining to attain the objectives, i.e., to identify the percentage of women who would be greatly benefitted by our app, and also to infer the percentage of users who face much difficulty in managing time. The detailed content and purpose of the tables are as stated below:

i Survey_General : It contains totally 11 fields as shown in table 1, from all the sectors in which working women face much impediments. Data mining is performed on it in order to find out the percentage of working women who will be greatly benefited by our app in order to make their daily miserable lives more abridged.

ii Survey_Time : It contains 5 fields related to time management, as shown in table 2. Data mining on this table is performed in order to identify the total percentage of women who are facing difficulty in managing their time between their professional and personal life on daily basis.

b) Data Mining through Clustering

Weka is a data mining open source tool to carry out various task like clustering, classification, association mining correlation and regression etc. Machine learning algorithm can also be easily developed and executed in Weka. We have used Weka tool to perform clustering on the gathered survey responses, so as to identify the percentage of working women who will be extremely benefited by our app.

Cluster is defined as a group of similar objects, all the objects of same group are related to each other in some manner viz. Characteristic. [7] The objects of one cluster are homogeneous in their group and are heterogeneous in another group. Clustering is an unsupervised learning process and not required the classes to be known in advance. It used the concept of learning by inspection or examination rather than by already stated examples.

General description about the clustering algorithm[8]:

1) Attributes should be normalized by considering highest and lowest values.
2) Assume the number of cluster and separate test data and actual data.
3) Calculate the distance of each sampled data from the cluster center.
4) A set of data is assigned to any one of the cluster with which it is less far.
5) The average of the entire columns known as centroid is calculated.
6) The distance of each sampled data can be calculated with the centroid. If its position fixed in the cluster and it will not shift to any other cluster, then its fine otherwise go back to step 3 and do all the process again up to step 6.

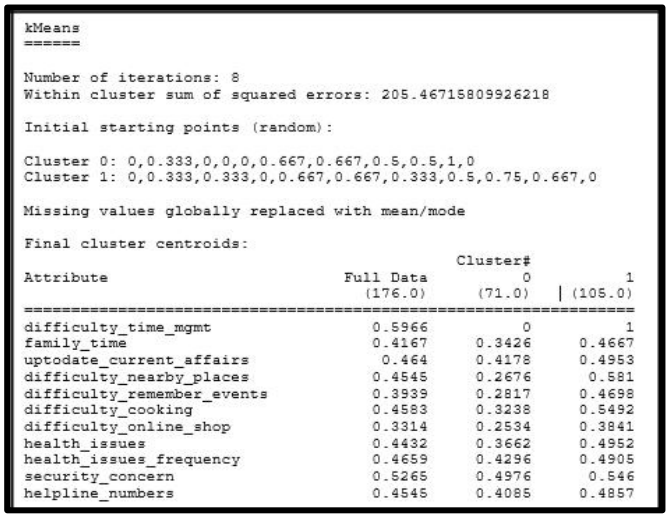
Clustering is an unsupervised learning method. K - Means is the simplest one among many available methods for clustering. Its procedure is very straightforward and effortless to categorize the data to some assumed set of K clusters. The basic theory behind this algorithm is to define K centers to adjust K clusters. K centers are located or placed in an intelligent manner so as it accommodate various data and produce separate results. For best result all the centers are placed far apart. Each data item are assigned to a nearest cluster or clusters with least distances.

Simple K-Means algorithm technique of data clustering has been performed on both the tables as it was observed to be the most efficient technique when compared with other data clustering algorithms in terms of highest data accuracy, lower error rate and least time consumed while running the algorithm.

c) Experimental Results

Data clustering technique has been performed on the tables, Survey_General and Survey_Time, derived from the data gathered from various working women across India. Data clustering has been adopted so as to perform unsupervised learning data mining technique.

The data clustering results of Simple K-Means clustering on Survey_General table is as follows in Figure 1:



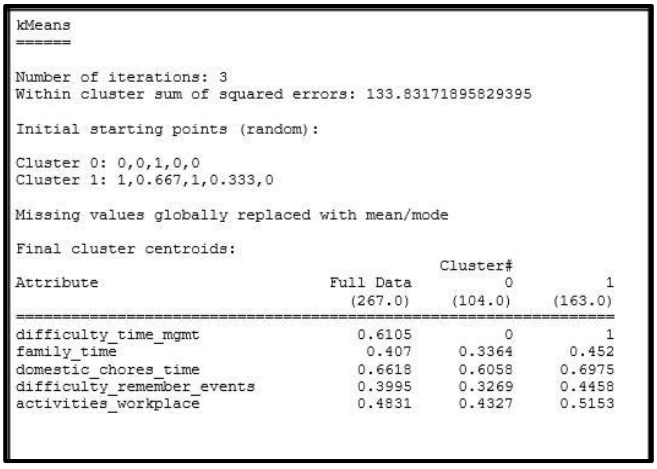
```
kMeans
======

Number of iterations: 8
Within cluster sum of squared errors: 205.46715809926218

Initial starting points (random):

Cluster 0: 0,0.333,0,0,0,0.667,0.667,0.5,0.5,1,0
Cluster 1: 0,0.333,0.333,0,0.667,0.667,0.333,0.5,0.75,0.667,0

Missing values globally replaced with mean/mode

Final cluster centroids:
                                        Cluster#
Attribute                 Full Data           0           1
                            (176.0)      (71.0)   | (105.0)
============================================================
difficulty_time_mgmt        0.5966           0           1
family_time                 0.4167      0.3426      0.4667
uptodate_current_affairs     0.464      0.4178      0.4953
difficulty_nearby_places    0.4545      0.2676       0.581
difficulty_remember_events  0.3939      0.2817      0.4698
difficulty_cooking          0.4583      0.3238      0.5492
difficulty_online_shop      0.3314      0.2534      0.3841
health_issues               0.4432      0.3662      0.4952
health_issues_frequency     0.4659      0.4296      0.4905
security_concern            0.5265      0.4976       0.546
helpline_numbers            0.4545      0.4085      0.4857
```

Fig. 1. Clustering result on Survey General Table

Here, Cluster 1 (60%) specifies the users who are highly in need of our app and will be using our app to manage their daily's frenzied life in a modest way and Cluster 0 (40%) implies the users who will be using less number of features of our app in order to manage their lives.

From the results of the clustering, we can infer that 60% of working women are facing hindrance in managing their day to day problems and hence, will be highly benefited by our application.

The screenshot of data clustering results of Simple K-Means clustering on Survey_Time table done in Weka is shown in figure 2:



```
kMeans
======

Number of iterations: 3
Within cluster sum of squared errors: 133.83171895829395

Initial starting points (random):

Cluster 0: 0,0,1,0,0
Cluster 1: 1,0.667,1,0.333,0

Missing values globally replaced with mean/mode

Final cluster centroids:
                                    Cluster#
Attribute                Full Data          0           1
                           (267.0)     (104.0)     (163.0)
============================================================
difficulty_time_mgmt       0.6105           0           1
family_time                 0.407      0.3364       0.452
domestic_chores_time       0.6618      0.6058      0.6975
difficulty_remember_events 0.3995      0.3269      0.4458
activities_workplace       0.4831      0.4327      0.5153
```

Fig. 2. Clustering result of Survey time table

Here, Cluster 1 (61%) implies the users who are facing much difficulty in managing time in their day to day lives, whereas Cluster 0 (39%) indicates the users who are facing less difficulty in time management.

From the results of the clustering, we can infer that 61% of working women are facing much difficulty in managing their day to day time for their personal as well as professional lives.

6. CONCLUSION AND FUTURE WORK

In this research work, we have presented numerous impediments faced by today's working women and gave the advantage of using android based application that rectifies all the issues of the extant systems. It is an amalgamated platform, with an easy to use interface that enables each individual to use it in an utmost convenient and competent manner.

- With the help of data mining and analysis techniques, we inferred that our app will contribute in enhancing daily lives of atleast 63% of working women, by resolving various snags faced by them in this real world.
- By applying K-means clustering algorithm, we also concluded that 61% of working women are facing much difficulty in managing their day to day time for their personal as well as professional lives.

In future, this work would be extended by taking the following things into consideration:

a) Stress level of the user will be calculated based on AI techniques and depending upon it, various suggestions and tips will be provided to the user.

b) A discussion forum will be included in the application so as to provide a common platform for working women from different corners to interact and get each other's problems solved.

c) The survey will be extended to get still more responses and perform different data mining techniques, so as to study the behaviour of several working women, as well as find still more difficulties faced by them in their daily lives and resolve them in our application.

REFERENCES

[1] Dileep Kumar S, "Health and Well Being of Working Women: A Socio-psychological Study", International Journal of Engineering Science andComputing, August 2016.

[2] G. Delina, R. Prabhakara Raya, "A study on Work-Life Balance in Working Women", IRACST – International Journal of Commerce, Business and Management (IJCBM), ISSN: 2319–2828 Vol. 2, No.5, October 2013

[3] Neetu Anand, Dhruvi Goyal, "Eazy_buzy life–An Android Application For Working Women", International Journal of Scientific Research In Computer Science, Engineering and Information Technology , 2456-3307, Volume 2 , Issue 6,2017.

[4] Neetu Anand, Dhruvi Goyal, "Design and Development Of Mobile App For Working Women –An Opportunity For Work-Life Balance" International Journal For Science And Advance Research In Technology, 2395-1052, Volume 3 Issue 12,2017.

[5] Neetu Anand, Dhruvi Goyal, "Working Woman's Problems and Challenges-A Survey", International Journal of scientific Research in Science and Technology, 2395-6011, Volume 3, Issue 8,2017.

[6] Raju, G.T. and Sathyanarayana, P. (2008). Knowledge discovery from Web Usage Data: Complete Preprocessing Methodology.IJCSNS, 2008.

[7] Kiruthika, M. and Dixit, D. (2010). Mining Access Patterns Using Clustering. International Journal of Computer Applications (0975 – 8887) Volume 4– No.11, August 2010.

[8] Kumar, P., Krishna, P.R., Bapi, R.S., De, S.K. (2007) Rough clustering of sequential data, Data and Knowledge Engineering 63 (2), (2007), 183–199.

TABLE 1: Survey General

S. No	Question	Variable	Nominal Value	Numeric Value
1.	Do you find it difficult to manage time between your personal and professional life?	difficulty_time_mgmt	Yes	1
			No	0
2.	Does after working hours, you get enough time for your family?	family_time	Not at all	1
			Sometimes	0.667
			Mostly	0.333
			Always	0
3.	Are you up-to-date in terms of current affairs going on?	uptodate_current_affairs	Not at all	1
			Sometimes	0.667
			Mostly	0.333
			Always	0
4.	Is it difficult for you to find nearby places like banks, hospitals, parks, beauty salon, etc.?	difficulty_nearby_places	Yes	1
			No	0
5.	Do you find it difficult to remember all your day to day important events like meetings, birthdays, PTMs, etc.?	difficulty_remember_events	Always	1
			Mostly	0.667
			Sometimes	0.333
			Not at all	0
6.	Do you find it difficult to cook nutritious quick and easy recipes?	difficulty_cooking	Always	1
			Mostly	0.667
			Sometimes	0.333
			Not at all	0
7.	Do you find it difficult to book cabs or shop online?	difficulty_online_shop	Always	1
			Mostly	0.667
			Sometimes	0.333
			Not at all	0
8.	Do you suffer from any health complications like mood swings,	health_issues	Yes	1
			Can't say	0.5

	depression, concentration problems due to problems at workplace?		No	0
9.	How often do you face health related problems?	health_issues_frequency	Most frequently	1
			Frequently	0.75
			Average	0.5
			Less Frequently	0.25
			Not at all	0
10.	Are you concerned about the security when you are out of home?	security_concern	Always	1
			Mostly	0.667
			Sometimes	0.333
			Not at all	0
11.	Do you remember women helpline numbers or is it saved in your cellphones?	helpline_numbers	No	1
			Yes	0

TABLE 2: Survey Time

S. No	Question	Variable	Nominal Value	Numeric Value
1.	Do you find it difficult to manage time between your personal and professional life?	difficulty_time_mgmt	Yes	1
			No	0
2.	Does after working hours, you get enough time for your family?	family_time	Not at all	1
			Sometimes	0.667
			Mostly	0.333
			Always	0
3.	How much time do you spend on domestic activities?	domestic_chores_time	Less than 2 hours	1
			2-4 hours	0.667
			4-6 hours	0.333
			More than 6 hours	0
4.	Do you find it difficult to remember all your day to day important events like meetings, birthdays, PTMs, etc.?	difficulty_remember_events	Always	1
			Mostly	0.667
			Sometimes	0.333
			Not at all	0
5.	Are you involved in any kind of activities in the workplace, community and society?	activities_workplace	No	1
			Yes	0