# Analysis and Study of the Scope of Improvements in Fitness Advisor Systems using Data Mining

**Christy Samuel Raju[1] Sanchit V Chavan[2] Karan Pithadia[3] Shraddha Sankhe[4] Prof. Sachin Gavhane[5]**
[1,2,3,4,5]Department of Information Technology
[1,2,3,4,5]Atharva College of Engineering, Malad(W).

*Abstract*— Today a fair number of health problems can be associated to a person's weight. Tackling this crisis with efficient diagnosis and spreading awareness about body weight related health problems is the need of the hour. The proposed system has a simple principle: "Better solution can be obtained only through better diagnostic methods". The proposed system is an application that advises the user to follow certain procedure to tackle his/her weight related problem. The system identifies three basic user requirements: gaining weight, losing weight and maintaining weight. The system uses data mining to efficiently direct all its users to the best possible solution. Clustering, Classification and Association algorithms are all used in the proposed system.

***Key words:*** Fitness, Advice, Clustering, Association, Classification, Data Mining, Health

## I. INTRODUCTION

The World Health Organization has made an estimation that obesity has doubled over the last 25 years. Many adults above 20 years of age were termed as overweight and 11% of them were found to be obese. Obesity boosts the risk of many disease.[1] Some of these diseases are excessively harmful to health.[6] If you are obese then you are at greater risk of developing severe health problems, that include heart disease, high blood pressure, gallstones, type 2 diabetes, breathing problems and certain cancers. That is why gaining and maintaining healthy weight is very important for overall health and can help you prevent and control various diseases and medical conditions. However there are also many with the reverse problem of being too skinny. This is another concern, because being underweight can also be just as bad for your health as being obese.

The proposed system adopts data mining as the tool to link the gap between the diagnosis solution and the user.[4] Today the prime diagnosis of weight problem is based on exclusively considering the BMI(Body Mass Index). The proposed system is an enhancement to this approach. The proposed system, considers various lifestyle factors (such as drinking, smoking, work hours, sleep etc) along with the BMI to effectively diagnose the weight related health problem.[3] Clusters are formed based on the said factors, a clustering algorithm like k-means is used to for the same. Later the clusters are associated by adopting association rules and further we apply classification and direct the users to various experts for advice.

## II. DETAILS EXPERIMENTAL

### *A. Proposed System*

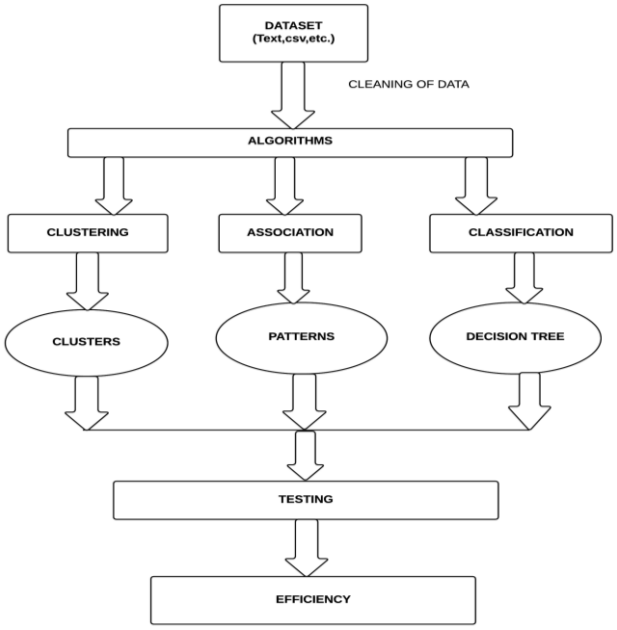


Fig. 1: System block diagram of the proposed system explaining the flow of procedures to be followed.

The proposed system can be divided in three different layers. In the initial layer the user enters all the needed data and a dataset is obtained. This layer also performs as the front end of the proposed system.

In the second layer the algorithms play a crucial role. The system includes clustering, association and classification algorithms.[3,4] This layer can be considered as the decision making layer of the proposed system as it directs the user to the expert advice. The clustering algorithm used in the proposed system is k-means and the obtained output are clusters. The clusters which are obtained are associated with each other using association patterns. Finally, classification is applied to these patterns and they are directed to the respective expert advice using a classification algorithm like ID3. The final layer is the testing part which is performed using a tool like WEKA. The confusion matrix which is obtained in WEKA is used check the efficiency and accuracy of the system.

## III. METHODOLOGY

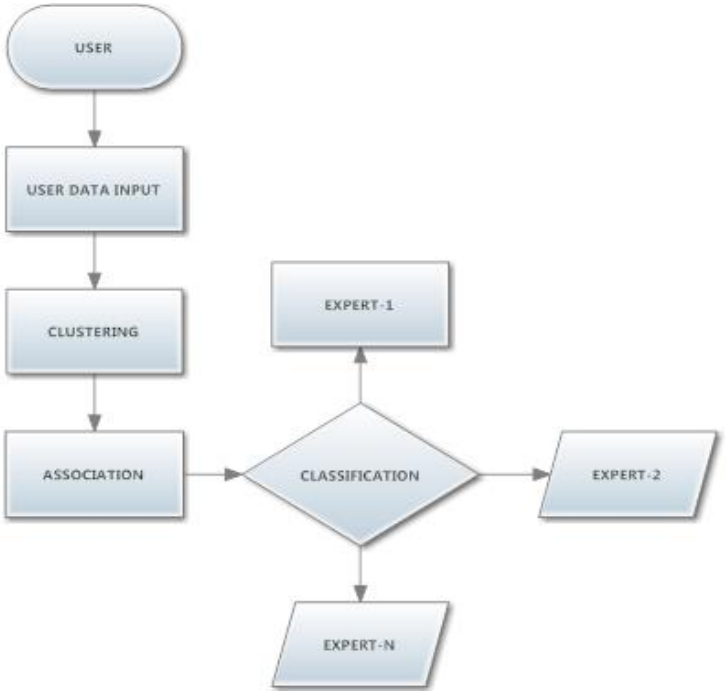


Fig. 2: Methodology of the proposed system illustrating steps involved in the entire process, displaying the flow of control and overall functioning of the system.

The proposed system effectively uses a combination of clustering, classification and association algorithms to efficiently deliver the best possible advice from the expert to the user's problem.

The functioning of the proposed system can be divided into four phases. In the initial phase, the user is asked to enter his requirement in the system i.e. the user mentions if he wants to lose, maintain or gain weight. Along with that, the user will also be asked to enter certain data(like weight, height, age, sex, body type etc) which will be utilised by the system to direct the user to the best solution to his/her problem. The system considers different important aspects of the user's daily lifestyle and makes sure that all these factors are incorporated in the decision making task.[3] These factors have been chosen based on the research done and keeping accuracy and efficiency in mind.

The factors considered by the system are as follows:

1) Weight
2) Height
3) Gender
4) Body Type
5) Drinking
6) Smoking
7) Medical Condition
8) Sleeping hours
9) Physical activity

BMI is calculated based on weight and height provided by the user. Once the data is entered, the user submits it.

In the next phase, the provided data is grouped to form different clusters based on the data provided by user using a clustering algorithm like k-means. Clusters are formed based on these sub-groups of the mentioned factors. These factors are divided into sub-groups as follows:

1) Body Type: Ectomorph, Mesomorph, Endomorph
2) BMI: Underweight, Normal, Overweight, Obese
3) Smoking: High, Medium, Low
4) Drinking: High, Medium, Low

5) Health Condition: Depends on the user
6) Sex: Male, Female
7) Sleeping hours: 5-5.99, 6-6.99, 7-7.99, 8-8.99, >=9
8) Physical Activity: High, Medium, Low

Physical activity: High, Medium or Low Based on the above stated sub-groups the user's data is accordingly grouped to form the respective clusters.

In the third phase, we will use association rules[4] on the clusters by which patterns are observed.

In the final phase, these patterns are finally directed to get the best expert advice for the respective pattern using a classification algorithm like ID3. The final output obtained from the system is a decision tree.

## IV. DESIGN DETAILS

### *A. E-R Diagram*

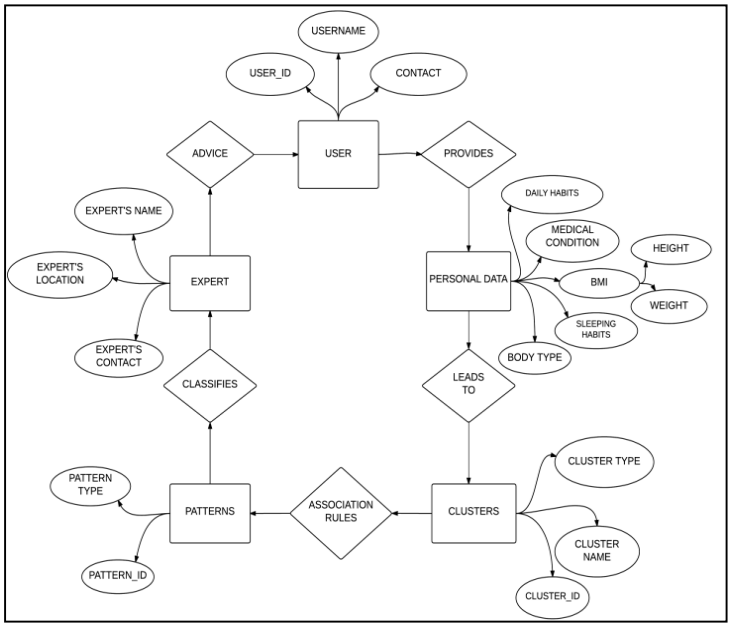


Fig. 3: E-R Diagram of the proposed system showing relationship between entities involved in the system.

### *B. Activity Diagram*

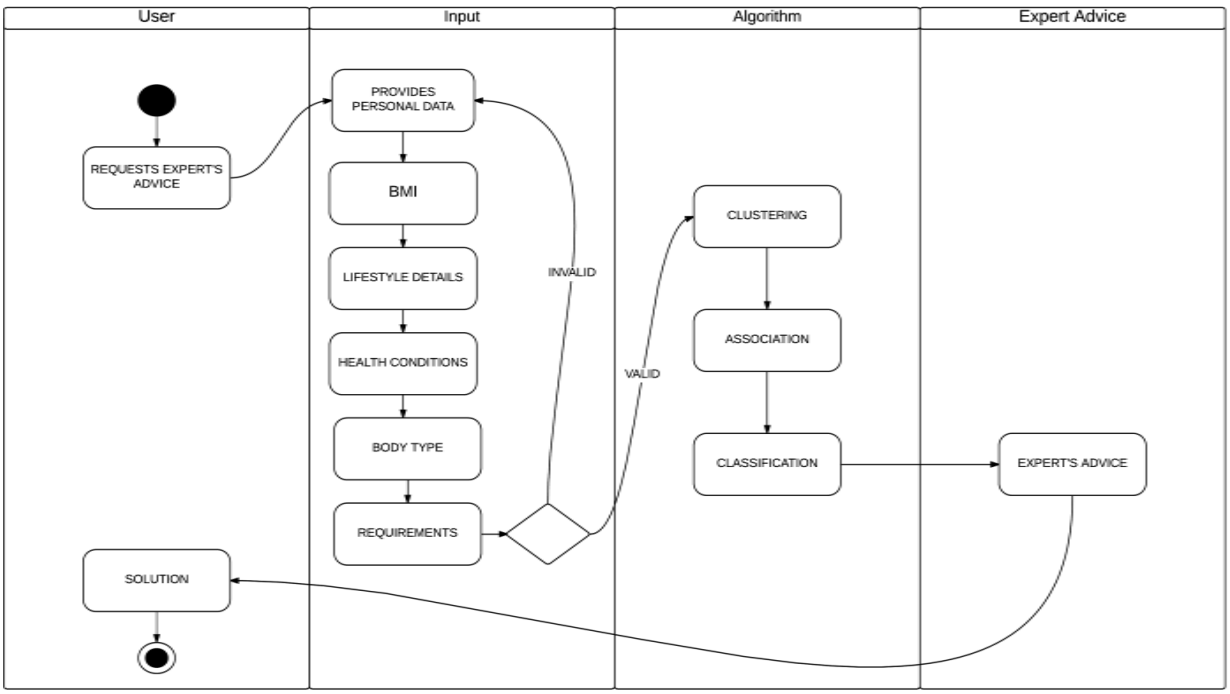


Fig. 4: Activity Diagram of the proposed system illustrating the order of activities followed in the system.

*C. Front End*

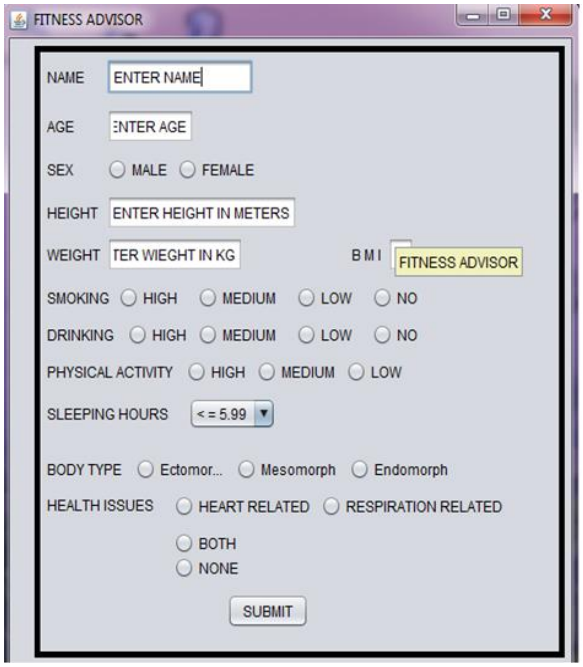


Fig. 1: Front End

## V. CONCLUSIONS

The following can be concluded from the above study.

1) BMI in itself is not accurate enough for diagnosing the weight related health problems hence other factors such as the user's lifestyle must be taken into account while diagnosis is done.
2) The proposed system can efficiently work as a cohesive data mining unit since the decision making process is automated.
3) If the number of experts is more it would directly lead to an increase in the number of advices available for the system to choose from for the user. This will train the system well making it less biased and thus improving the efficiency of the system which in turn will lead to a better diagnosis and better advice.
4) The system can be deployed at any health centre and can be used to direct any customer to the best expert available for tackling the customer's problem.

## ACKNOWLEDGMENTS

We would like to deeply express our sincere gratitude to Project coordinators. We are eager and glad to express our gratitude to the Head of the Information Technology Dept. Prof. Neelima Pathak, for her approval of this project. We are also thankful to her for providing us the needed assistance, detailed suggestions and also encouragement to do the project. We would like to deeply express our sincere gratitude to our respected principal Prof. Dr. Shrikant Kallurkar and the management of Atharva College of Engineering.

## REFERENCES

[1] Fletcher Lu, Kei Turner, Bernadette Murphy, " Reducing Adolescent Obesity with a Mobile Fitness Application: Study Results of Youth Age 15 to 17", DOI: 10.1109/HealthCom.2013.6720738 Conference: 2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013) .
[2] Fabio Buttussi, Luca Chittaro, "Smarter Phones for healthier lifestyles: An adaptive Fitness Game", IEEE Pervasive Computing, vol.9, no. 4, pp. 51-57, October-December 2010, doi:10.1109/MPRV.2010.52 .
[3] Muhamad Hariz B. Muhamad Adnan, Wahidah Husain, Nur`AiniAbdul Rashid, "Parameter Identification and Selection for Childhood Obesity Prediction Using Data Mining "School of Computer Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia
[4] Irina Tudor, "Association Rule Mining as a Data Mining Technique", BULETINUL universitatii Petrol-Gaze din Ploiesti, Vol. LX No1/ 2008.
[5] Mehbod Tavallaee, Ebrahim Bagheri, Wei Lu,AliA Ghorbani , "A Detailed Analysis of the KDD Cup 99 Data Set"Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on 8-10 July 2009.
[6] Er. Bindu Rani, Er. Parneet Kaur , Er. Avinash Bansal , "Analysis of Student Physical Fitness Data Using Data Mining Algorithm",International Journal of Applied Engineering Research, ISSN 09734562Vol.7 No.11 (2012).