RESEARCH DESIGN ON STUDY OF WORK AND HEALTH LIFE BALANCE WITH WORKING WOMEN USING DATA MINING AND WEBSERIES

Sarikonda Vishali[1], *Dr. Ramakrishna Seemakurti[2]
[1] Research scholar, computer science, JJTU, Rajasthan, India, 333001(Asst Professor, SVIT, Secunderabad-500 003)
[2] Principal, Ph.D(Pricipal, SVIT, Secunderabd, India, 500003) (Approved Research Guide, JJTU, Jhunjhunu-333001, Rajasthan, India)

***Abstract*:**This paper proposes the effective ways to provide health life balance for working women with the help of data mining and webseries. Due to the increase in economic conditions and social demands, the necessity for the women working is increasing day by day.As a result working women have tremendous pressure to develop a career while upholding active engagement in their personal life.But they are getting very less time for themselves due to the work pressure.Hence , it is very essential for the working women to get a balance between work and personal life. The increasing responsibilities on the personal front with the technological blessings like advanced mobile phones, notepads, etc. that keeps work life integrated with personal life also creates stress on personal and professional fronts in this knowledge age. This affects the person's physical, emotional and social well-being. Thus, achieving work life balance is a necessity for working women to have a good quality of life. This paper is an attempt to explore the tough challenges faced by working women in maintaining a balance between their personal and professional life and also to provide health life balance with the help of data mining.

Keywords: *Health Life Balance, Working Women, Data Mining(DM), CBR (Case Based Reasoning), RBR(Rule Based Reasoning), IDSS(Integrated Decision Support Systems)*

I. INTRODUCTION

Work life balance mainly describes the balance between an individual's personal life and professional life. A healthy work-life balance assumes great significance for working women particularly in the current context in which both, the family and the workplace have posed several challenges and problems for women. The dynamics of the work environment have exerted enormous pressure on working women as they need to cope with virtually two full time jobs – one at the office and the other at home .So the necessity for the balance between the work and family.

Women of the early centuries were mostly confined to their kitchens and those who were employed worked in factories, farms or shop works. Very few women had the access to higher education and they were forced to be at the mercy of their fathers' or husbands' attitudes towards women and work. The fast developing knowledge economy has given place for more number of women to be enlightened by higher education. Education has not only empowered them but also has given them robust careers. With brain power being the requisite skill in this knowledge era, rather than endurance or physical strength, the women workers seem to flood into every industry on par with men. But this has indeed become a tough challenge for women as they have to perform a lot of duties in home and office as well.

Working mothers of today fulfill family responsibilities and also try to remain fully involved in their careers coping up with the competing demands of their multiple roles. The caring responsibilities that working mothers have lays a heavy stress on them when it is combined with their professional duties. The attempt of working women to integrate, organize and balance the various problems and activities in their different roles simultaneously puts themunder tremendous pressure.

II. LITERATURE REVIEW

Data mining in health care has become increasingly popular because it offers benefits to care providers, patients, healthcare organizations, researchers, and insurers.

Care providers can use data analysis to identify effective treatments and best practices. By comparing causes, symptoms, treatments, and their adverse effects, data mining can analyze which courses of action are most effective for specific patient groups. It can also identify clinical best practices to help develop guidelines and standards of care.Patients can receive better, more affordable healthcare services. This is especially true when healthcare managers use data mining applications to identify and track chronic diseases and high-risk patients, design appropriate interventions, and reduce the number of hospital admissions and claims.

Data Mining tools and principles are mainly used in health care expert systems to extract different relationships with in the data elements as knowledge. By combining different data mining techniques with expert systems, a new system known as "Integrated Decision Support System" (IDSS) is proposed, to get the better results. It converts knowledge into desired format and uses different tools are used for the construction of its architecture. To reduce possible solutions for diabetic diagnosis, Rule Based Reasoning (RBR), Case Base Reasoning (CBR), and Web Based Portal Joint Asia Diabetes Evaluation(JADE) programs are integrated with Reliable Access and Probablistic Inference based on clinical Data (RAPID) in the developed IDSS system to enhance existing systems for fast extraction of knowledge.All the disease(like Heart Disease, Cancer, HIV) are the leading cause of death all over the world. Some of the diagnostic and laboratory procedures are invasive,costly and painful to patients. Single Data Mining Technique in the diagnosis of all disease has been comprehensively investigated showing acceptable levels of accuracy. Still, using data mining techniques to identify a suitable treatment for all disease has received less attention.

The data is related to the Minimum data Set (MDS) with respect to the mental health patients. The MDS-MH is a standard assessment tool for evaluating patients having problems related to mental health. There are a number of MDS tools are available for the mental assessment.

One of their advantages is that they are cross applicable to the other forms of MDS databases, so that with this knowledge we can apply the same tools to all the different types of data that are obtained and expect to get the same outcome in terms of accuracy.

III. PROPOSED METHOD

3.1 Data Mining

A Data is said to be a fact expressed in some metrics. The relevant, reliable, processed & organized data is called Data Base, which can be summarized in 'Data ware House'. 'Informatica' a tool is used to store data by integrating, transforming for consistent view& query from different sources.With the time & increase in operations, the amount of data is grown /generated at very fast in business & scientific communities. While assets, sales turnover, payroll, area and customer data comes from business, tera bytes of results come from science experiments. Similarly, an experiment produces tera bytes of data in science and technology.

An oft metaphor say "we are drowning in data& yet starving for knowledge" sums up the current situation. We are rich of data, but poor of knowledge in accurate decision making.

Data Mining(DM)is a process of extraction through intriguing (relevant, constructive, unexplored & valuable) and finding "patterns" or clusters from large amounts of data using tools, algorithms, machine learning and statistics. It is a knowledge discovery (KD) process for developing a model of quick visualization. It explores the links, relationships& associations which are hidden in large data base. It is more of an art than a science, because of irreplaceable preconception& intuition of knowledge that human mind possess in decision making. Tools like Weka, kneel, R..etc are used. With the emergence of intranet(LAN,MAN,WAN) and internet(World Wide Web), the integration of heterogeneous data created large amounts of data in repositories with less than rigid structure. The boundaries among the countries got erased to share data across many fields & platforms. The web has become a big repository of data.

Initially, a unified frame work for expressing, optimizing and processing data mining problems over relational data is carried out. Next, defining the problem of data mining structure followed from the semi structured data. Verification Model & Discovery Model is prepared based on hypothesis & emphasis.

Objectives of data mining:

- To apply association rule mining to electronic medical records (EMR) to discover sets of risk factors and their corresponding subpopulations & use of summarizing techniques for easy clinical use.
- To evaluate data mining tools in medical and health care operations.
- To apply association rule mining & represent patients at particularly high risk of developing a disease & cluster them based on symptoms & laboratory results.
- To extract useful knowledge through diagnosis, therapy and prognosis.
- To design query blocks & system architecture based on past data using data mining techniques.
- To conduct a comparative evaluation & provide guidance regarding applicability, strengths & weaknesses.
- To propose extensions to incorporate risk of any disease into the process of finding an optimal summary.
- To apply Bottom-Up Summarization (BUS) algorithm to produce suitable summary.
- To improve useful knowledge for individuals, organizations and public health departments to take preventive measures through awareness about trend with knowledge, but not to practice medicine.

IV. DISCUSSIONS

Fuzzy C Means Clustering:

Fuzzy C-Means clustering is an improvement over HCM clustering. In FCM-clustering each data point belongs to a cluster to a degree specified by a membership grade and it allows one piece of data to belong to two or more clusters. The membership matrix U is allowed to have elements with values between 0 and 1, while, the sum of degrees of belongingness of the data point to all clusters is always equal to unity:

$$\sum_{i=1}^{c} u_{ij} = 1$$

The cost function for FCM is generalization of

$$J=\sum_{i=1}^{c} J_i =\sum_{i=1}^{c}\sum_{j=1}^{n} u^{m}_{ij}\ d^{2}_{ij}$$

where uij is between 0 and 1, ci is the cluster center of fuzzy group i; ij i j d = c -x is the Euclidean distance between the ith cluster center and the jth data point; and m Є [1,∞]. The necessary conditions for (9) to reach to its minimum are

$$c_{i=} \frac{\sum_{j=1}^{n} u_{ij}^{m} X_j}{\sum_{j=1}^{n} u_{ij}^{m}}$$

and

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left[\left\|\frac{c_i - X_j}{c_k - X_j}\right\|\right]^{2/(m-1)}}$$

The FCM Algorithm [20] works iteratively through the preceding two conditions until no more improvement is noticed. In a batch mode operation, FCM determines the cluster centers c_i and the membership matrix U using the following steps:

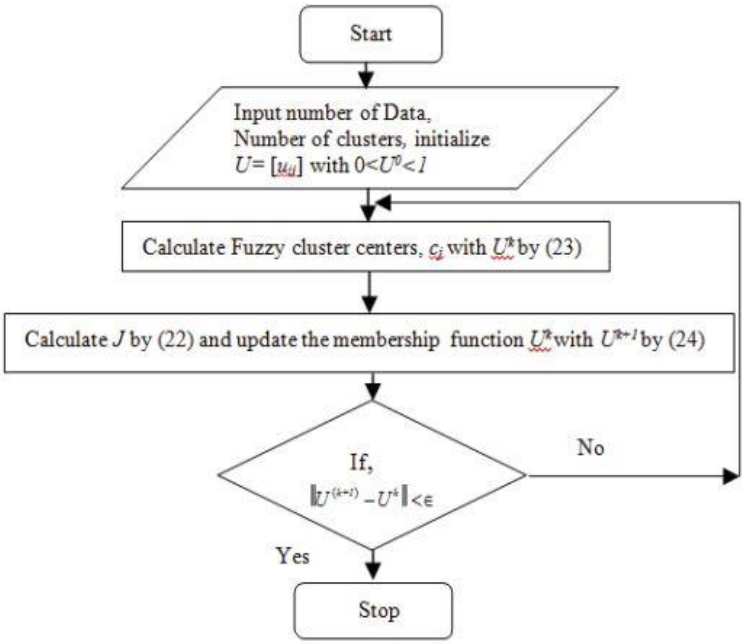
Step 1. Initialize the membership matrix U with random values between 0 and 1 such that the constraints are satisfied.

Step 2. Calculate c fuzzy cluster centers c_i, i=1,…,c

Step 3. Compute the cost function. Stop if it is below a certain tolerance value or its improvement over previous iteration is below a certain threshold.

Step 4. Compute a new U and go to step 2.

FCM algorithm starts by assigning random values to the membership matrix U, therefore several iterations have to be conducted for getting good performance and it is possible to have different degree of membership function to each cluster. The algorithm must be run several times, each starting with different values of membership grades of data points.



Advantages

1) Gives best result for overlapped data set and comparatively better then k-means algorithm.
2) Unlike k-means where data point must exclusively belong to one cluster center here data point is assigned membership to each cluster center as a result of which data point may belong to more then one cluster center.

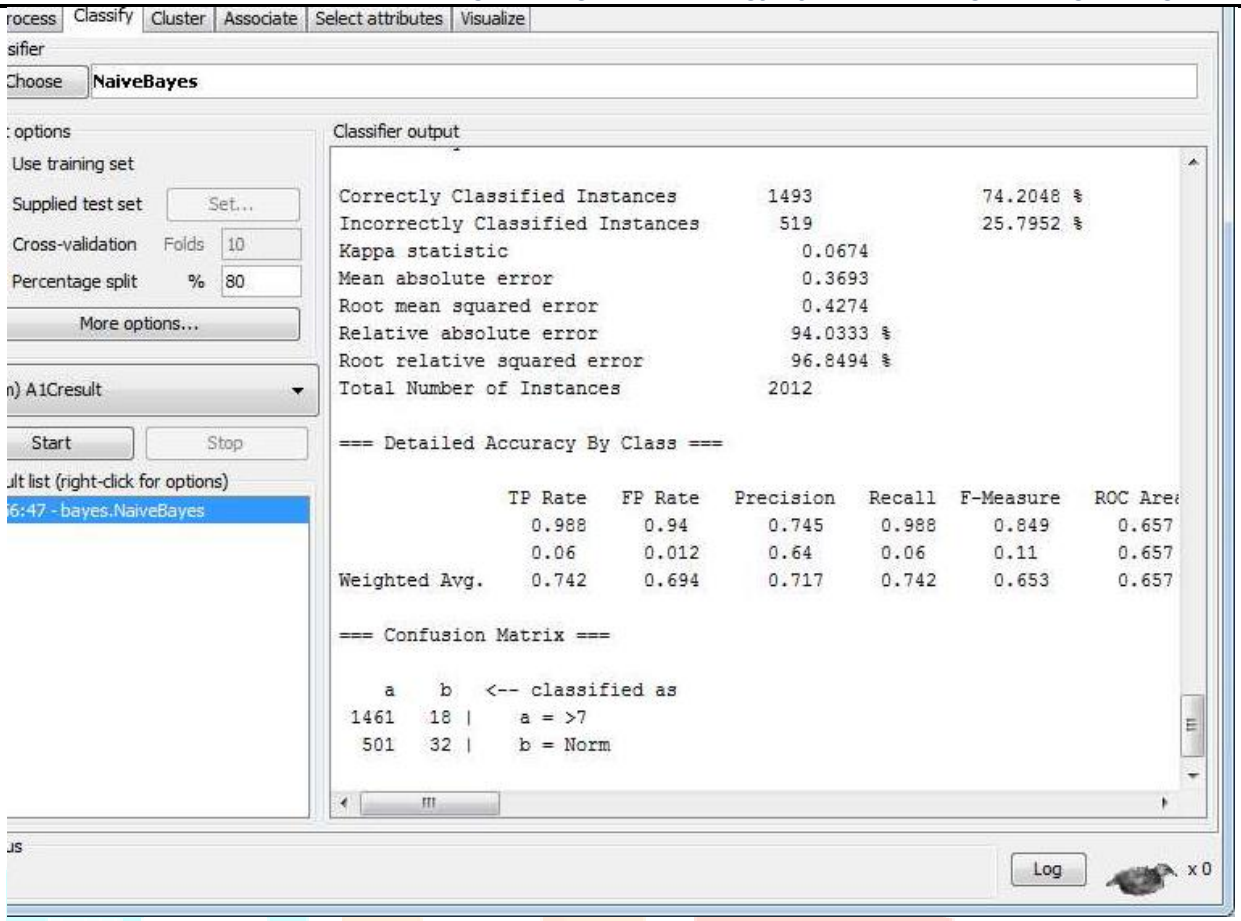
Disadvantages

1) Apriori specification of the number of clusters.
2) With lower value of β we get the better result but at the expense of more number of iteration.

V. RESULT ANALYSIS

NAIVE BAYES ALGORITHM:

This experiment, Logistic algorithm was used and three sizes to train the model were selected 50%, 65% and 80%. Figure presents the model that used 80% variations.Generally, the analysis of the evaluation with the confusion matrix, which illustrate the amount of correctly and incorrectly classification classes, the correct classifications denoted by (TP) True Positive and True Negative (TN). The incorrect predicted outcomes occur when a false positive (FP) appear, that is when an outcome is predicted as positive (yes) while it is actually negative (no), furthermore, the same happened when a false negative (FN) occur if an outcome is predicted as negative but it is actually positive.[24] It is significant to use F-Measure, it produces a high result when Precision and Recall are both balanced, although Precision and Recall are valid measures, but the F-Measure was used because, Precision and Recall can



VI. CONCLUSION

With dual career couples widely prevalent in this modern era, there is a need for the balance between the work life and personal life. But more research is neede Considerably more research is needed to gain additional insight into the work–family balance. This paper deals with usage of data mining techniques to balance thr health and work.This technique was able to measure women's work–life balance and found weekly hours of work and the stress associated with work were very important determinants of women's work–life balance, alongside their occupations, age and caring responsibilities. Conflicts in work-life balance of working women affects their health who report more stress, headaches, muscle tension, weight gain and depress than their male counterparts. Juggling between the obligations towards the families and expectations of the organisation and constant struggle to maintain a balance between work and family can have serious implications on the life of an individual by affecting their well-being and overall quality of life. There is a widespread demand from womens for the right to balance work and home life in today's busy world where finding time for oneself seems impossible. Health and wellness programs can, for sure help working women in balancing their personal and professional life. But they alone cannot be the answer to addressing the problems of imbalance.The problems and difficulties of women are multi-dimensional as evident from the literature reviewed; therefore, they require further probing to help working women in balancing their work and family life.

REFERENCES

[1] Bailyn, Lotte; Drago, Robert; and Kochan, Thomas A.; "Integrating Work and Family Life – A Holistic Approach". A Report of the Sloan Work-Family Policy Network. 9/14/2001: pp 1- 10.
[2] R. Baral & S. Bhargava; "HR interventions for work-life balance: evidences from organisations in India". International Journal of Business, Management and Social Sciences, Vol. 2, No. 1, 2011, pp. 33-42
[3] Bharat, S. 2003. "Women, work, and family in urban India, Towards new families?" in J. W. Berry, R. C. Mishra, and R. C. Tripathi ed., Psychology in human and social development, Lessons from diverse cultures pp.155-169 New Delhi, India, Sage.
[4] Clark SC (2000).Work/Family Border Theory: A NewTheory of Work/Family Balance. Human Relations 53(6): 747–770.
[5] Daniels, L. and McCarraher, L. (2000) The Work–Life Manual. London: Industrial Society.
[6] Denise Horner Mitnick, "The Impact of Working Women on Work/Life Balance Perspectives", University of Pennsylvania ScholarlyCommons, 2006
[7] Francene Sussner Rodgers, "When the Business Case Is Common Sense: Coming to Terms with America's Family Challenge," ACA Journal, (Autumn 1992)
[8] Friedman, S D and Greenhaus, J H (2000). "Work and family—Allies or enemies? What happens when business professionals confront life choices", New York: Oxford University Press.
[9] Grossman, Allyson Sherman. "Working Mothers and Their Children". Monthly Labor Review: Special Labor Force Reports – Summaries. May, 1981: pp 49 -54.
[10] Jeffrey H. Greenhaus, Karen M. Collins & Jason D. Shaw, "The relation between work–family balance and quality of life", Journal of Vocational Behavior 63 (2003) 510–531.
[11] Khalid Hammouda, Fakhareddine Karray, "A Comparative Study of Data Clustering Techniques, University of Waterloo, Ontario, Canada.

[12] Jesus Gonzalez, Ignacio Rojas, Hector Pomares, Julio Ortega, & Alberto Prieto, "A New Clustering Technique for Function Approximation" IEEE Transaction on Neural Networks, Vol.13, NO.1, January, 2002.

[13] Ricardo, Bellazzia & Blaz, Zupan (2008) "Predictive data mining in clinical medicine: current issues and guidelines", International journal of medical informatics. vol. 77. 8197.

[14] Adam, Kiezun, I-Ting, Angelina, Lee & Noam, Shomron, (2009) "Evaluation of optimization techniques for variable selection in logistic regression applied to diagnosis of myocardial infarction", Bioinformation. Vol.3. 311-313.