International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Diabetes Mellitus Recommendation System using Association Rule Mining

Sudhanshu S. Deshpande[1], Swati S. Joshi[2]

[1]M.E (Computer) Department of Computer Science & Engineering, N.B.Navale Sinhgad College of Engineering, Solapur 413255

[2]M.E (Computer) Assistant Professor, Department of Computer Science & Engineering, N. B. Navale Sinhgad College of Engineering ,Solapur 413255

Abstract: ***Early detection of the diabetes mellitus with an elevated risk of developing it, is critical to take overall clinical care of patients. Diabetes leads to critical medical diseases including ischemic heart disease, stroke. Early identification of patients who is at risk of developing diabetes is a major healthcare need for today's world. This study aims to apply a association rule mining to EMR (Electronic Medical Records) to discover the risk of developing a diabetes. Association rule mining generates a large set of rules which needs to summarize and use for recommendation purpose. The Subpopulations found after this summary covers peoples having at high risk of developing diabetes.***

Keywords: Diabetes, Association rule, mining, EMR, medical

1. Introduction

The WHO (World Health Organization) estimates that by 2030 there will be an approximate 350 million peoples with having type2 diabetes, Associated with many diseases for example heart problems, strokes, hypertension, finding the person who has developing diabetes or at a high risk of developing it is a challenge.

As far as medical concern, Diabetes is part of the metabolic syndrome, which is having a constellation of diseases including hyperlipidemia, hypertension and central obesity , These diseases are in fact interacting with each other and they are related with each other , with cardiac and vascular diseases and thus understanding of these and modelling these interactions is very important to understand.

Association rules are implications that associate a set of potentially interacting conditions with elevated risk of diabetes. The use of association rules is particularly beneficial in these type of cases, because in addition to quantifying the diabetes risk for a person, they also readily provide the physician with kind of sureness, namely the associated set of conditions in this case. This set of conditions can be used to guide the treatment towards a more personalized, sophisticated and targeted preventive care or diabetes management for suffering peoples. The resulting rule sets from this can sometimes be very large. Especially, in this proposed work, we considered a rich set of different risk factors, namely laboratory results, medications which we used and information which is commonly available EMR systems.

Key contributions are as follows.
We are presenting a clinical application of association rule mining to identify the different sets of co-morbid conditions that imply significantly increased risk of diabetes.

2. Related Works

1) "G. S Collins, S. Mallett, O. Omar, and L.-M. Yu, "**Developing risk prediction models for type 2 diabetes**: A systematic review of methodology and reporting"", they conducted a search of PubMed and EMBASE databases to find out studies which describes the development of models combining two or more variables to predict the risk of prevalent or incident type 2 diabetes.
2) In "Wilson Soto and Amparo Olaya–Benavides, "**A Genetic Algorithm for Discovery of Association Rules**", A genetic algorithm is proposed in this article for discovery of association rules. The main characteristics of the algorithm are: (1) The individual is represented as a set of rules (2) The fitness function is a criteria combination to evaluate the rule's quality – high precision prediction, comprehensibility and interestingness – (3) Subset Size-Oriented Common feature Crossover Operator (SSOCF) is used in the crossover stage (4) mutation is calculated through non-symmetric probability and selection strategy through tournament and (5) the algorithm was implemented using the library *lambdaj*.
3) In "R. Agrawal and R. Srikant, "**Fast algorithms for mining association rules**," They consider the problem of discovering association rules between items in a large database of sales transactions. They present two new algorithms for solving this problem that are fundamentally different from the known algorithms. Empirical evaluation shows that these algorithms outperform the known algorithms by factors ranging from three for small problems to more than an order of magnitude for large problems.

3. System Architecture

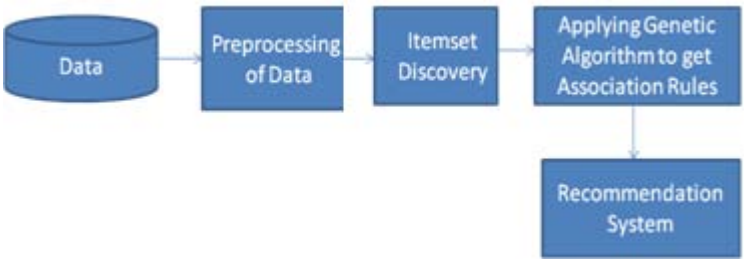

Figure 1: System Architecture

- Gather medical data.
- Preprocess medical data to remove noisy data.
- Convert Filtered data into usable data.
- Compute Frequent Itemset.
- Genetic Algorithm for rule finding
- Initialize Genetic Algorithm – Selection, Fitness function , Selection, Crossover, Mutation
- Calculate fitness of chromosomes if (Fitness > Threshold) then add to rule.

4. Conclusion

This study aims to apply association rule mining to EMR to discover sets of risk factors that represent patients at particularly high risk of developing diabetes. By Using Genetic algorithm to discover the Association Rules, Which gives an optimized solution, which need for easy clinical use. This Study is going to evaluate this technique on real world prediabetic patient cohort. This system recommends about the risk of Diabetes Mellitus.

References

[1] G. S Collins, S. Mallett, O. Omar, and L.-M. Yu, "Developing risk prediction models for type 2 diabetes: A systematic review of methodology and reporting," *BMC Med.*, 9:103, Sept. 2011.
[2] Wilson Soto and Amparo Olaya–Benavides, "A Genetic Algorithm for Discovery of Association Rules", 2011 30th International Conference of the Chilean Computer Science Society
[3] H. S. Kim, A. M. Shin, M. K. Kim, and N. Kim, "Comorbidity study on type 2 diabetes mellitus using data mining," *Korean J. Intern. Med.*, vol. 27, no. 2, pp. 197–202, Jun. 2012.
[4] F. Afrati, A. Gionis, and H. Mannila, "Approximating a collection of frequent sets," in *Proc. ACM Int. Conf. KDD*, Washington, DC, USA, 2004.
[5] J. Tuomilehto *et al.,* "Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impared glucose tolerance," *N. Engl. J. Med.*, vol. 344, no. 18, pp. 1343–1350, May 2001.
[6] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th VLDB*, Santiago, Chile, 1994.