

# Association Rule Summarization for Relative Risk Analysis of Diabetes Mellitus

Ranjan R. Sorte<sup>1</sup>, Charudatta B. Bante<sup>1</sup>, Aditya P. Thakare<sup>1</sup>, Shubham T. Saha<sup>1</sup>, Mosam B. Meshram<sup>1</sup>,  
Prof. Vishesh P. Gaikwad<sup>2</sup>

<sup>1</sup>BE Students, Department of Computer Science and Engineering Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

<sup>2</sup>Assistant Professor, Department of Computer Science and Engineering Priyadarshini Bhagwati College of Engineering, Nagpur, Maharashtra, India

## ABSTRACT

Diabetes is a creating pandemic of non-transmittable disease which impacts most of the overall public on the planet. Remembering the ultimate objective to smother the advancement of diabetes mellitus we use association control summary to electronic therapeutic records to discover set of peril factors and the contrasting sub-people which talks with patients at particularly high risk of making diabetes. Regularly alliance control mining makes enormous volume of educational accumulations which we need to diagram for any restorative record or any clinical use. We join four methodologies to find the fundamental segments which incite high threat of diabetes all these four procedures made rundowns that delineated sub masses at high risk of diabetes with each methodology having its unmistakable quality. According to our inspiration we use bottom up summarization (BUS) estimation which conveys more proper once-over.

**Keywords:** Diabetes Mellitus, Data mining, Association Rule Mining, Survival Analysis, Association Rule Summarization

## I. INTRODUCTION

Diabetes mellitus is creating disease ailment which impacts more than 25.8 million people and around 7 million of them don't know they have this ailment. Ordinarily diabetes is a get-together of infirmities depicted by high glucose (blood glucose). Exactly when a man has diabetes the body either makes enough insulin or not ready to use its own specific insulin effectively. Right when glucose gets create in our blood, that glucose should be controlled or ought to be enough used else it may to lead passing. The risk of death of a man who has diabetes is twice as the person who does not have diabetes of same age.

The genuine ensnarement of diabetes are coronary sickness and stroke. Adults with diabetes have coronary sickness end rates around 2 to 4 times

higher than adults without diabetes the peril of stroke is 2 to 4 times higher among people with diabetes. It moreover prompts hypertension and 67% of diabetic patients have circulatory strain more vital than or equal to 140/90 millimetres of mercury or used professionally endorsed remedy for hypertension. Diabetes is a principle wellspring of visual inadequacy among adults developed 20-74 years. Around 60% 70% of people with diabetes have mellowed to extraordinary kinds of tactile framework hurt. The delayed consequence of such mischief consolidate crippled sensation or misery in the feet or hands showed preparing of sustenance in the stomach carpal entry issue or other nerve issue.

Ideal around 30% of people with diabetes developed 40 years or more prepared have obstructed sensation in the feet. Diabetes may moreover provoke

unpredictability in the midst of pregnancy, insufficiently controlled diabetes before beginning and in the midst of the vital trimesters of pregnancy among women with sort 1.

Diabetes can achieve noteworthy birth surrenders in 5% to 10% of pregnancy and unconstrained untimely births in 15% to 20% of pregnancies. On other hand for a women earlier diabetes upgrading blood glucose levels previously and in the midst of early pregnancy can be reduce the risk of birth disfigurements in their infant kids. Incapably controlled diabetes in the midst of the second and third trimesters of pregnancy can achieve preposterously huge babies speaking to a risk to both mother and child.

Association chooses are recommendation that accomplice set conceivably collaborating conditions (e.g.: high BMI and the proximity of hypertension conclusion). The usage of alliance fundamentals is particularly helpful in light of the fact that despite estimating the diabetes peril, the similarly quickly outfit the specialist with a "legitimization" specifically the related course of action of conditions. These conditions can be used to direct treatment towards a more redid and concentrated on preventive care or diabetes organization.

Diabetes is a bit of the metabolic issue, which is a great assortment of ailments including hyperlipidaemia (lifted triglyceride and low HDL levels), (hypertension) and central weight (with weight record outperforming 30 kg/m<sup>2</sup>). These diseases speak with each other, with cardiovascular and vascular sicknesses and in this way perception and showing these associations is basic. Alliance norms are proposals that accomplice a course of action of possibly working together conditions (e.g. high BMI and the closeness of hypertension analysis) with lifted risk. The usage of association benchmarks is particularly favourable, in light of the fact that despite assessing the diabetes shot, they also instantly give the specialist a "guard", particularly the related

course of action of conditions. This course of action of conditions can be used to control treatment towards a more redid and concentrated on preventive care or diabetes organization. While connection rules themselves can be adequately deciphered, the resulting oversee sets would some be able to of the time be considerable, crumbling the interpret limit of the lead set when all is said in done. Particularly, in this work, we think about a rich strategy of hazard fragments, particularly co-troubling ailments, lab results, solutions and estimation data that are typically open in electronic medicinal record (EMR) frameworks. With such a wide arrangement of hazard factors, the strategy of found standards winds up observably combinatorial broad, to a size that incredibly prevents translation. To beat this test, we related supervise set summarization systems to pack the chief lead set into a more traditionalist set that can be deciphered effectively. Diverse reasonable association speak to set digest procedures have been proposed [10] yet no unmistakable course exists with respect to the real nature, attributes and shortcomings of these technique. The centralization of this interesting copy is to audit and depict four existing association lead rundown systems and give direction to specialists in picking the most sensible one. A regular insufficiency of these systems is their failure to consider diabetes risk— a steady outcome. Recollecting a definitive target to influence these procedures more too proper, we anticipated that would incidentally transform them: we extend them to combine data about interminable result factors.

In particular, our key obligations are as per the going with.

- We show a clinical utilization of union lead mining to perceive sets of co-dispiriting conditions (and the patient sub people groups who experience the malicious effects of these conditions) that prescribe outright broadened risk of diabetes.

- Association control mining on this wide strategy of segments accomplished an exponentially broad game-plan of association measures. We made four regular association run set diagram systems (on a very basic level from the survey [10]) by consolidating the danger of diabetes into the way toward finding a flawless once-finished.
- Our fundamental obligation is a relative assessment of these improved structure systems that gives direction to experts in picking a fitting figuring for a for all intents and purposes indistinguishable issue.

## II. RELATED WORKS

A diabetes list is in a general sense an adroit model that allocates a score to a patient in context of his surveyed hazard of diabetes. Collins et al. [7] facilitated a wide review of diabetes records depicting the hazard factors and the demonstrating structure that these archives used. They found that most records were incorporated substance in nature and none of the illustrated archives have considered coordinated effort among the danger segments.

While we don't consider any new diabetes archive passed on after the outline, a present review [12] concentrating on the metabolic issue (of which diabetes is a segment) addresses an essential progress. Kim et al. utilized organization together run mining to deliberately analyze co-occasions of conclusion codes. The resulting organization together rules don't constitute a diabetes record in light of the way that the study does not delegate a specific eventual outcome of intrigue and they don't survey or suspect the danger of diabetes in patients, in any case they found some tremendous association between finding codes.

We have beginning late attempted a diabetes consider [4] where we expected that would find the associations among pains in the metabolic issue. We utilized a vague companion from this present overview; notwithstanding, we included just eight finding codes and age as pointers. We found partnership rules including some of these eight finding codes, explored the hazard of diabetes that these measures give on patients and displayed the standards as an advancement diagram sketching out how patients advance from a solid state towards diabetes. We demonstrated that the approach discovered clinically crucial association presumes that are steady with our accommodating need.

With just eight pointer factors, the level of the found direct set was modest– 13 imperative rules– and thusly, clarification was quick. In actuality, no oversee set summary was focal.

## III. LITERATURE SURVEY

Chaudhari et al [13] Disease affirmation is a champion among the most essential livelihoods of such framework as it is one of the guideline wellsprings of going wherever all through the world. Foresee the human utilize the responsibilities from complex tests facilitated in labs also expect the illness considering hazard parts, for example, tobacco smoking, liquor affirmation, age, family history, diabetes, hypertension, raised cholesterol, physical lethargy, weight. Agents have been utilizing two or three data mining techniques to help remedial organizations authorities in the examination of coronary ailment. K-Nearest-Neighbor (KNN) is one of the feasible data mining procedures utilized as a bit of demand issues. Beginning late, agents are demonstrating that joining specific classifiers through voting is overcoming other single classifiers. This paper asks about applying KNN to help human organizations pros in the complete of ailment outstandingly coronary illness. It also asks about if arranging voting with KNN can upgrade its exactness

in the affirmation of coronary sickness patients. The outcomes display that applying KNN could accomplish higher exactness than neural system gathering in the finding of coronary sickness patients. The outcomes moreover demonstrate that applying voting couldn't upgrade the KNN accuracy in the confirmation of coronary ailment.

Prof. Mythili et al [12] Diabetes mellitus, in fundamental terms called as diabetes, is a metabolic ailment, where a man is influenced with high blood glucose level. Diabetes is a metabolic issue expedited because of the slip-up of body to make insulin or to appropriately use insulin. This condition rises when the body does not make enough insulin, or in light of the way that the cells don't react to the insulin that is passed on. Blood glucose test is the basic framework for diagnosing diabetes. Additionally, there have been diverse robotized procedures proposed for complete of diabetes.

Every single one of these systems has a few data esteems which would be the aftereffect of various tests that ought to be done in recovering core interests. This paper proposes a structure those game plans to support the patients experiencing assorted medicinal tests, which by a long shot the greater part of them consider as a dull endeavor and redundant.

The parameters perceived for diagnosing diabetes have been made in a way that, the client can expect on the off chance that he is affected with diabetes himself. Back Propagation tally is utilized for conclusion.

Ahmed et al [15] Heart ailment is a basic reason behind dreariness and mortality in bleeding edge society. Remedial conclusion is essential yet caught endeavor that ought to be performed unequivocally and enough. The able data examination instruments are utilized to expel pleasing picking up from the monstrous measure of helpful data. There is tremendous data open inside the remedial

organizations structures. Regardless, there is an undertaking of persuading examination devices to find canvassed associations and cases in data. Learning introduction and data mining have discovered different applications in business and trial space.

One of the applications is infection finding where data mining instruments are showing beneficial results. This examination paper proposed to discover the heart issue through data mining, Support Vector Machine (SVM), Genetic Algorithm, unpleasant set hypothesis, association rules and Neural Networks. In this survey, we promptly evaluated that out of the above frameworks Decision tree and SVM is best for the coronary disorder. So it is watched that, the data mining could help in the perceiving check or the want of high or all things considered safe heart issue.

Thangaraju et al [16] Data mining is the showing of looking prior databases with a specific genuine goal to make new data. There are distinctive sorts of data mining procedures are accessible. Strategy, Clustering, Association Rule and Neural Network are likely the hugest structures in data mining. In Health mind associations, Data mining acknowledge a basic part. Most an incredible piece of the time the data mining is utilized as a bit of human organizations endeavours for the course toward predicting ailments. Diabetes is an unending condition. This proposes it proceeds for quite a while, reliably for some person's entire life [11]. This paper considers the examination of diabetes gaging approaches utilizing gathering structures. Here we are utilizing three unique sorts of clumping systems named as Hierarchical social affair; Density based assembling, and Simple K-Means gathering. Weka is utilized as a device.

Durairaj et al [17] Neural Networks are one of the touchy enrolling systems that can be utilized to make wants on remedial data. Neural Networks are known as the Universal pointers. Diabetes mellitus or

basically diabetes is a pollution accomplished because of the augmentation level of blood glucose.

Particular customary methodology, considering physical and innovation tests, are accessible for diagnosing diabetes. The Artificial Neural Networks (ANNs) based framework can adequately relate for hypertension risk want. This enhanced model isolates the dataset into both of the two get-togethers. The prior disclosure utilizing touchy enlisting strategies help the experts to diminish the likelihood of stopping all tomfoolery of the malady. The data set decided for ask for and exploratory excitement depends on upon Pima Indian Diabetic Set from (UCI) Repository of Machine Learning databases. In this paper, an isolated review is driven on the use of various delicate planning structures for the want of diabetes. This overview is depended upon to see and propose a productive strategy for prior gauge of the ailment.

#### IV. IMPLEMENTATION

We try to utilize association rule mining to the electronic supportive record (EMR); All the hazard learn about a patient particularly co-disheartening sickness and research office results and game plans are being accessible in the EMR, there are less opportunities to miss encounters about a patient with the far reaching blueprint of threat factors the game-plan of watched danger winds up being inconceivably tremendous to vanquish this we utilize rule set design methodology which is utilized to pack the primary rule set into a traditionalist set. We utilize the running with systems:

1. APRX-collection
2. RPG-global
3. TOPK
4. BUS.

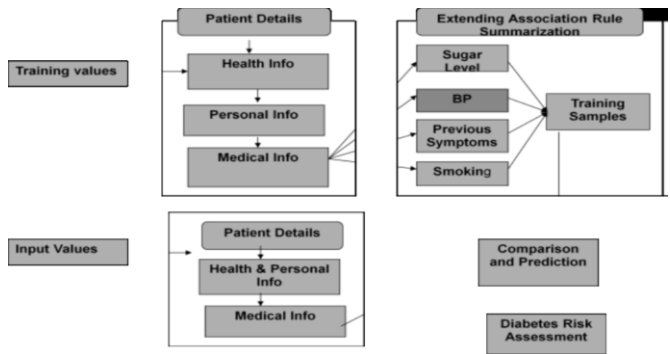


Figure 1. System Architecture

We first check the assistance of individual things and comprehend which of them wide (ie.) we have least help are. In each pass we begin with a seed set of things saw to clear in the past pass. We utilize this seed set for making new possibly liberal thing sets called contender key, thing set and number the real reinforce for these hopeful thing set amidst the indifference the data.

At the entire of the pass we understand which sure thing set are really expansive and they progress toward getting the chance to be seed for not pass. This technique contains until the point that no new impressive thing set are found. Testing quantifiable criticalness: for each found thing set we need to test whether the result disseminating in the affected and unaffected subpopulation is unmistakably specific.

Step-2 the game-plan of thing set is secluded with the objective that lone the unquestionably essential ones are returned as distributional association rule, this rule is portrayed by the running with bits of information from the measure of thing set gathered. Let OR be the watched number of diabetes scene in the subpopulation DR secured by R. enable ER to mean the common number of diabetes occasions in the subpopulation secured by R.

$ER = OR - i \in DRy_i$  where  $y_i$  is the martingale for patient.

The relative risk factor is defined by R

$$RR = OR/ER.$$

**Table 1.** Description of the risk factors that appeared in any of the summarized rules

Parameter	Weightage	Values
Male & Female	Age<30 >30to<50	0.1 0.3 0.7 0.8
Smoking	Never Past Current	0.1 0.3 0.6
Overweight	Yes No	0.8 0.1
Alcohol intake	Never Past Current	0.1 0.3 0.6
Heart rate	Low(<60 bpm) Normal(60 to 100bpm) High(>100bpm)	0.9 0.1 0.9
Blood sugar	High(>120&<400) Normal(>90&<120) ) Low(<90)	0.5 0.1 0.4
Bad cholesterol	Very high>200 High(160 to 200) Normal<160	0.9 0.8 0.1

When we try to apply distributional rule mining with our electronic helpful records it influenced a sweeping to number of (quantifiably essential) rules. Rules that were made conceivably move from each other inciting jumbling of clinical patters. With a specific genuine goal to vanquish the issue of this wide number of rules which were conveyed we go for gathering the rule set into more humble set for our less asking for review. We first investigation the present rule set and database outline frameworks then we try to merge a nonexclusive structure with a specific genuine target to get a relentless outcome of variable into record.

Eventually we exhibit the rule set conveyed by the broadened format estimation, for every figuring we utilized the parameter setting that gave the best outcomes to APRX Collection we utilized  $\alpha = 0.1$ ,  $\lambda = 1$  for RPG general we utilized  $\delta = 0.5$ ,  $\sigma = 0.2$ ,  $\lambda = 0.98$  for top K we utilized  $\lambda = 0.2$  and for BUS we utilized  $\lambda = 0.1$

**A. APRX-collection**

The APRX Collection estimation is utilized to discover the supersets of the condition (hazard consider) in the rule with a definitive target that most subsets of synopsis rule will be honest to goodness rules in the primary (un merged) set and these subset rules propose for all intents and purposes indistinguishable hazard for diabetes.

**Table 2.** Rule set summarized by APRX- collection

R	RR	ER	OR	RULE
1	1.96	36.24	71	Fibra
20	1.34	271.7 1	36 3	Bmi trigal acerab Statin aspirin htn
16	1.19	426.7 8	50 6	Hdl trigl acearb Aspirin htn
15	1.31	348.9 2	45 7	Bmi trigal statin aspirin ihd
10	1.23	534.5 8	66 0	Bmi sbp ccb htn

The APRX gathering revolves just around clarification of the rule along these lines it needs data about which patients are beginning at now secured in like manner patients can get secured by different rules actuating rules with in a general sense the same as condition this method additionally needs in precision and data about high risk subgroups.

**B. RPG-global**

The rule disadvantages of APRX Collection were the plenitude in the rule set and the incapacitating of the threat. The RP-Global outline looks like APRX Collection n in that it is by and large worried with the surge of the rule and in this manner it plays out a to an awesome degree serious weight. RPG Global has two disadvantages by contemplating Patient augmentation and by building the format from rules in the essential rule set.

**Table 3.** Rule set created by RP-Global.

RR	ER	OR	RULE
1.69	32	55	Bmi trigal acearb diuret htn
1.23	52	65	Acearb bb diuret aspirin htn
1.29	42	55	Sbp tchol acearb diuret htn
2.10	25	54	Hdl trigal diuret aspirin htn
1.28	42	54	Bmi tchol hdl trigl tobacco

**C. TOPK**

Top-K algorithm decreases the emphasis in the rule set which was conceivable through tackling patients as opposed to the attestation of the rules. This approach surrendered the outstanding weight rates of past two count TOP-K still accomplishes high weight rate and it satisfactorily perceived rules with high hazard and low excess.

**Table 4.** Rule created by the top-k algorithm

RR	ER	OR	RULE
2.40	21.70	52	Fibra htn
1.58	37.97	60	Bmi hdl ihd
1.47	45.52	67	Sbp htn tobacoo
1.46	317.03	464	Bmi htn
1.62	32.16	52	Sbp tchol trigl statin htn

**D. BUS**

The designs made by BUS (appeared in Table 5) and Top-K are equivalent in quality. The BUS graph exhibits less changeability (it tends to utilize tantamount conditions: bmi and trigl co-happen in 40% of the rules), yet this decreased instability does not change over into broadened plenitude in the patient space.

Transport (instead of Top-K) manages the patients and not on the rules. In like way, excess to the degree rule verbalization can happen. Regardless, BUS unequivocally controls the wealth in the patient space through the parameter charging the base number of new (starting at now revealed) cases (patients with diabetes occasion) that should be secured by each rule. Along these lines the diminished fluctuation in the rule verbalization does not change over into broadened emphasis.

**Table 5.** Top 10 summarized rule created by BUS

RR	ER	OR	RULE
2.34	24	57	Bmi trigal acearb statin htn
2.10	25	54	Hdl trigal diuret aspirin htn
1.91	56	107	Bmi trigal statin htn
1.54	78	121	Bmi trigal tobacco
1.37	39	54	Dbp diuret htn

The BUS rule set understands how to meld coronary ailment prior (rule #3) and with higher risk (2.15) than the Top-K rule set. In like way, BUS incorporates tobacco use in a blend of threat factors with higher relative hazard than Top-K. As a rule, paying little personality to the refinements in the computations, BUS and Top-K make for all intents and purposes indistinguishable remarkable once-overs.

**V. EXPERIMENTAL RESULT**

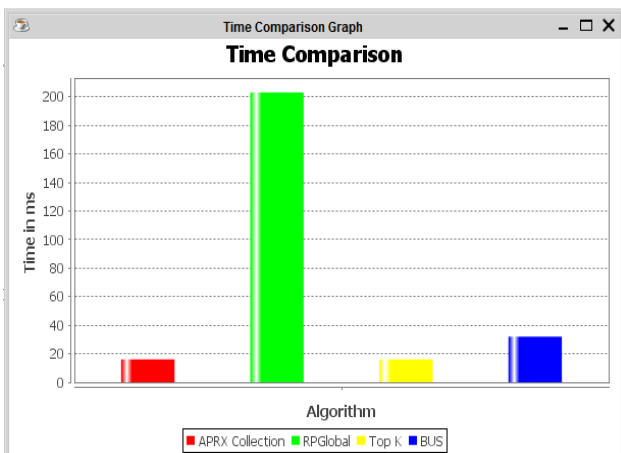
The amount of rules ought to have been diminished to a level where clinical clarification is feasible. To this end, we focused four procedures to pack these rules into sets of 10-20 rules that clinical analysts can evaluate.

While each one of the four procedures made sensible summaries, each system had its unmistakable quality.

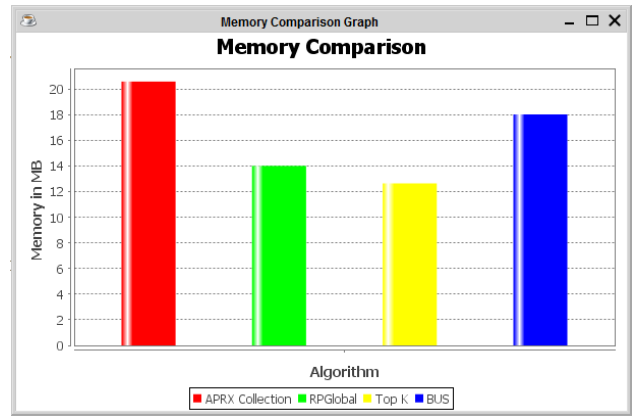
Regardless, not these characteristics are basically profitable to our application. We found that the most basic differentiator between the calculations is whether they use an assurance model to join a rule in the framework in perspective of the surge of the rule or in light of the patient sub masses that the rule covers.

APRX-COLLECTION and RP-Global fundamentally chip away at the surge of the rules with a basic focus of extending weight. They use appoint rules, each of which addresses different special rules. Such illustrative rules achieve high weight, however debilitate the risk of diabetes over the frequently gigantic subpopulation they cover. Top-K and BUS work fundamentally on the patients and their objective– particularly in case of Top-K– can be considered as restricting overabundance. They conveyed awesome layouts in light of the way that a supportive response of decreasing redundancy is to achieve incredible weight. The inverse isn't legitimate: high weight rate does not realize low reiteration.

Among Top-K and BUS, we found that BUS held possibly more redundancy than Top-K, which allowed it to have better patient degree and better ability to revamp the main data base. This ideal position made BUS the most suitable calculation for our inspiration.



**Figure 2.** Time Comparison



**Figure 3.** Memory Comparison

Our Result also shows the Time and space utilization of the above four algorithms. Figure show the results generated for the same.

## VI. CONCLUSIONS

The electronic data created by the usage of EMRs in routine clinical practice can support the disclosure of new learning. Alliance control mining coupled to an abstract framework gives a fundamental gadget to clinical research. It can uncover covered clinical associations and can propose new cases of conditions to redirect neutralizing activity, organization, and treatment approaches. In our specific representation, we used distributional connection control mining to recognize sets of peril components and the relating understanding subpopulations that are at through and through extended risk of progressing to diabetes. An over the best quantities of association rules were discovered thwarting the clinical explanation of the results.

## VII. REFERENCES

- [1]. F. Afrati, A. Gionis, and H. Mannila, "Approximating a collection of frequent sets," in Proc. ACM Int. Conf. KDD, Washington, DC, USA, 2004.
- [2]. R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in Proc. 20th VLDB, Santiago, Chile, 1994.
- [3]. Y. Aumann and Y. Lindell, "A statistical theory for quantitative association rules," in Proc. 5th KDD, New York, NY, USA, 1999.



- [4]. P. J. Caraballo, M. R. Castro, S. S. Cha, P. W. Li, and G. J. Simon, "Use of association rule mining to assess diabetes risk in patients with impaired fasting glucose," in Proc. AMIA Annu. Symp., 2011.
- [5]. Centers for Disease Control and Prevention. "National diabetes fact sheet: National estimates and general information on diabetes and prediabetes in the United States," U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2011 [Online].
- [6]. V. Chandola and V. Kumar, "Summarization-Compressing data into an informative representation," *Knowl. Inform. Syst.*, vol. 12, no. 3, pp. 355–378, 2006.
- [7]. 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.
- [8]. Diabetes Prevention Program Research Group, "Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin," *N. Engl. J. Med.*, vol. 346, no. 6, pp. 393–403, Feb. 2002.
- [9]. G. Fang et al., "High-order SNP combinations associated with complex diseases: Efficient discovery, statistical power and functional interactions," *PLoS ONE*, vol. 7, no. 4, Article e33531, 2012.
- [10]. M. A. Hasan, "Summarization in pattern mining," in *Encyclopedia of Data Warehousing and Mining*, 2nd ed. Hershey, PA, USA: Information Science Reference, 2008.
- [11]. Rakesh Motka, Viral Parmar, Balbindra Kumar, A. R. Verma, "Diabetes Mellitus Forecast Using Different Data Mining Techniques", International conference on computer and Communication Technology
- [12]. Prof.Sumathy, Prof.Mythili, Dr.Praveen Kumar, Jishnujit T M, K Ranjith Kumar, "Diagnosis of Diabetes Mellitus based on Risk Factors", International Journal of Computer Applications, Vol.10, Issue No.4, November.2010
- [13]. Anand A. Chaudhari, Prof.S.P.Akarte, " Fuzzy and Data Mining based Disease Prediction using K-NN Algorithm", International Journal of Innovations in Engineering and Technology, Vol. 3, Issue No. 4, April 2014
- [14]. Prof.Sumathy, Prof.Mythili, Dr.Praveen Kumar, Jishnujit T M, K Ranjith Kumar, " Diagnosis of Diabetes Mellitus based on Risk Factors", International Journal of Computer Applications, Vol. 10, Issue No. 4, November 2010
- [15]. Aqueel Ahmed, Shaikh Abdul Hannan, " Data Mining Techniques to Find Out Heart Diseases: An Overview", International Journal of Innovative Technology and Exploring Engineering, Vol. 1, Issue No. 4, September 2012
- [16]. P. Thangaraju, B.Deepa, T.Karthikeyan, "Comparison of Data mining Techniques for Forecasting Diabetes Mellitus", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue No. 8, August 2014
- [17]. M. Durairaj, G. Kalaiselvi, " Prediction Of Diabetes Using Soft Computing Techniques- A Survey", International Journal of Scientific & Technology Research, Vol. 4, Issue No.3, March 2015
- [18]. S.F.B, Jaafar and Darmawaty Mohd Ali. "Diabetes Mellitus Forecast using Artificial Neural Network (ANN), Asian conference on sensors and the international conference on new techniques in pharmaceutical and medical research proceedings (IEEE), Kuala Lumpur, Malaysia, 5-7 September 2005, pp 135-139.
- [19]. S. Alby, B. L. Shivakumar," A survey on data-mining technologies for prediction and diagnosis of diabetes", International conference of IEEE 2014.
- [20]. Gyorgy J. Simon, Terry M. Therneau, Steven S. Cha, " Extending association rule summarization techniques to assess risk of diabetes mellitus", IEEE VOL 27, no. 1, January 2015