



Risk Factor Prediction Of Cardiovascular Problems For Diabetic Mellitus Patients Using Rule Based Classification

¹S.Oviya, ²Ms. M. Meena Krithika, M.Sc.,M.Phil.,(Ph.D)

¹M.Sc. Computer Science Student,²Assistant Professor

¹PG Department of Computer Science,

¹Nallamuthu Goundar Mahalingam College, Pollachi, Tamilnadu

ABSTRACT :Different epidemiological studies have shown that type 2 diabetes is associated with a 2-to-3-fold increased risk for sudden cardiac death. This system proposes a new scheme to detect the possibilities for the cardiac arrest based on the diabetic patients socioeconomic profile and their activities. The algorithm analyses to find a subset of instances that are the best fit and based on that is able to predict the outcome. Initially user should register in this application with basic personnel details. After completion of successful registration, user should login into this application with registered user name and password. Login authenticated users will get the home page to enter their health related details to predicting the risk factor level for cardiac arrest or heart related problems. Age, Gender, Smoking habit, Obesity, Alcohol intake, High salt diet, High saturated fat diet, Physical activity with exercise detail, Sedentary life style, Bad cholesterol, Blood pressure, Blood sugar and Heart rate are the required data to predict the risk factor for the cardiac arrest using rule mining. This project is to be developed using Python with Scikit-learn as front end and MySQL as back end. Scikit-learn machine learning library for Python.

Keywords :Diabetic to Cardiac, Rule based classification, Risk factor

1. BACKGROUND

Diabetes mellitus is a group of metabolic diseases in which a person has high sugar count in the blood. This high blood sugar produces the symptoms of frequent urination, increased thirst, increased hunger and dehydration. Untreated, diabetes can cause many complications in the body and it makes difficult to treat other diseases. Serious long-term complications include heart disease, kidney failure, and damage to the eyes. The existing systems are available to predict the diabetic mellitus and cardiac problems based on personal and socio economic profiles. Naïve Bayes classifier, Support Vector Machine and Decision tree based classifiers are used to predict the diabetic mellitus. Type 1 , Type2 and Gestational diabetics possibilities and risk factors are predicted using data mining classification techniques. Naïve Bayes, KNN, SVM and Reptree based classifiers are used to predict the cardiac problems. But influences of diabetic to Cardiac arrest risk factors are not implemented with high accuracy level of prediction. Naïve Bayes classifier is used to predict the association between diabetic and cardiac problem. Limitations in the previous works are follows

- Diabetic mellitus based risk factors and cardiac problem based risk factors only predicted
- Diabetic mellitus influences in cardiac problems are not predicted
- Association between cardiac and influences are not determined

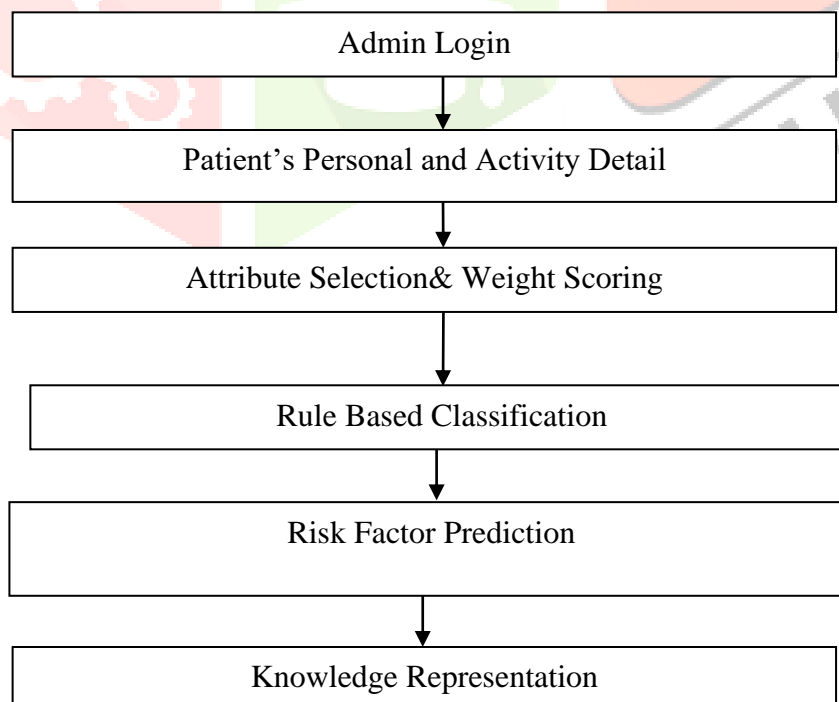
2. NEED OF PROPOSED WORK

Disease prediction plays an important role in data mining. Diagnosis of a disease requires the performance of a number of tests on the patient. However, use of data mining techniques, can reduce the number of tests. This reduced test set plays an important role in time and performance. This proposed system helps to predict the diabetic influences and risk factors for the diabetic patients to face the cardiac problems. This system proposes a new scheme to detect the possibilities for the cardiac arrest based on the diabetic patients socioeconomic profile and their activities using rule based classifier with weightage scoring. The advantages of the of the proposed system as follows

- Determines the association between diabetic and cardiac problems
- Predicts diabetic influences in cardiac problems with Risk factors
- Accuracy in prediction
- Weightage scores are taken from medical repository to increase the risk factor prediction accuracy

3. PROPOSED TECHNIQUE

The reduction of mortality from sudden cardiac arrest (SCA) within the putting of coronary heart sickness (CHD) stays a prime venture, specially amongst patients with kind 2 diabetes. Diabetes is associated with an accelerated threat of SCA, as a minimum in part, from an accelerated presence and volume of coronary atherosclerosis (macrovascular disorder). Diabetes is also associated with microvascular ailment and autonomic neuropathy; and, these non-coronary atherosclerotic pathophysiologic procedures additionally have the capacity to growth the chance of SCA. Diabetes is intently associated with a set of interrelated determinants of SCA hazard, which include: 1) the presence and extent of atherosclerotic coronary artery disease that results in clinical CHD; 2) patchy areas of myocardial fibrosis that outcomes in impaired left ventricular (LV) filling (diastolic disorder) and/or LV systolic coronary heart failure; and, three) abnormalities of electrical propagation in the myocardium that is meditated in electrocardiographic repolarization and conduction abnormalities, e.G. Prolongation of the QT—interval and QRS—c language. Those characteristics are regarded to be related to an increased chance of SCA within the network.



This task goals to are expecting the danger factors for the diabetic patients to affect the cardiac arrest or cardiac primarily based disease by way of analysing the diabetic affected person's profile thru facts mining classification method. It enables to product the risk factors with greater accuracy than different prediction fashions.

Modules of the assignment as follows

- Admin Login
- Person Profile
 - Private details
 - Way of life and pastime info
 - Idl cholesterol, Sugar and BP degrees
- Choice Tree and Rule primarily based type
- Weightage Scoring
- Threat element Prediction
 - Rule based totally Prediction
 - Weightage Scoring Prediction

Admin Login

Administrator is the authorised person of this prediction machine. To start with administrator ought to login into this machine using person name and password. After completion successful login manner, admin gets the main menu to maintain the user profile, perform the prediction method using classification and view the anticipated effects.

Consumer Profile

User profiles are maintained by using the administrator with affected person's fundamental non-public details, working and living style, meals and other habits. Patient profile includes affected person identification, call of the patient, region of residing, marital popularity, gender, age in years, smoking dependancy, obese, salt eating regimen, high saturated fats weight loss plan, exercise habit, period in hours, diabetic in familial, terrible cholesterol level, blood pressure level, blood sugar stage, coronary heart price, menopauses popularity (in case of gender is woman). From all attributes of the consumer profile, 14 attributes are decided on to carry out the class and prediction technique. Personal information like call, marital popularity, location of living, etc., attributes are ignored. Critical 14 attributes are recognized and selected for the procedure.

Rule based totally class

In rule based totally category, the discovered version is represented as a fixed of IF-THEN policies. The policies are composed of components namely rule antecedent - that is the If element, and rule consequent - that's the else part. An IF-THEN rule is of the form IF situation THEN conclusion. The situation in the rule of thumb antecedent has attribute assessments and the ensuing contains class prediction. As an instance bear in mind the guideline R1

R1: excessive CHD hazard ← gender = girl ^ age > 63 ^ frame mass index >25 kg/m²

This rule describes a collection of sufferers who're lady, overweight, and older than 63 years as having high chance of Coronary heart ailment (CHD). The policies may be derived from caused choice bushes. To extract regulations from a decision tree, one rule is created for each direction from the foundation to a leaf node. Each splitting criterion is logically ANDed to form the rule of thumb antecedent and the leaf node has elegance prediction that is the rule consequent. Rules can also be extracted immediately from statistics. Here one rule is learned at a time. Whilst a rule is found out, the statistics tuples blanketed by the rule of thumb are eliminated and the method repeats on the closing facts tuples. The policies accordingly generated can then be used to categorise a new information tuple based totally on the rule of thumb caused with the aid of the facts tuple.

Weightage Scoring

Weightage scoring process is based on the age range of the user, gender, smoking habit, obesity, alcohol intake, diet with salt and saturated fats, regular exercise, sedentary life style, level of bad cholesterol, blood pressure and sugar and heart beat rate. Each parameter have weightage score based on the data range. Weightage score is applied to the parameters; calculate the sum and average of weightage score. Average value is used to predict the risk factor.

Parameters	Weightage	
Male and Female	Age < 30	0.1
	>30 to <50	0.3
	Age>50 and Age <70	0.7
	Age>70	0.8
Smoking	Never	0.1
	Past	0.3
	Current	0.6
Overweight	Yes	0.8
	No	0.1
Alcohol Intake	Never	0.1
	Past	0.3
	Current	0.6
High salt diet	Yes	0.9
	No	0.1
High saturated fat diet	Yes	0.9
	No	0.1
Exercise	Never	0.6
	Regular	0.1
	High If age < 30	0.1
	High If age > 50	0.6
Sedentary Lifestyle/inactivity	Yes	0.7
	No	0.1
Hereditary	Yes	0.7
	No	0.1
Bad cholesterol	Very High >200	0.9
	High 160 to 200	0.8
	Normal <160	0.1
Blood Pressure	Normal (130/89)	0.1
	Low (< 119/79)	0.8
	High (>200/160)	0.9
Blood sugar	High (>120&<400)	0.5
	Normal (>90&<120)	0.1
	Low (<90)	0.4
Heart Rate	Low (< 60bpm)	0.9
	Normal (60 to 100)	0.1
	High (>100bpm)	0.9

Danger component Prediction

Threat issue is an occasion, situation or characteristic that is found in a topic, that is commonplace in sufferers of a specific sickness. On this assignment, prediction of chance aspect to have an effect on the cardiac arrest or cardiac primarily based disease for the sufferers who're suffered with diabetic mellitus problems. Due to diabetic mellitus particularly influenced one for cardiac troubles.

Excessive threat, Medium danger and occasional hazard are the elements recognized by this project using two methodologies which include Rule based totally Prediction and Weightage Scoring Prediction. Rule primarily based prediction is based at the generated regulations and Weightage scoring is based at the score values assigned to the attributes based totally on the situations.

4.CONCLUSION

Health care associated records are large in nature and they come from numerous birthplaces which aren't appropriate in shape or high-quality. In recent times, the usage of know-how and revel in of copious specialists and scientific screening facts of patients gathered in a database for the duration of the prognosis manner, has been extensively time-honored. The challenge may be utilized in hospitals and research centers to research the primary fitness related attributes and how they make a contribution to the diabetic with cardiac arrest. Essentially diabetics and cardiac associated problems are interrelated one. Records mining and its rule mining has incredible significance for location of medicine, and it represents comprehensive manner that needs thorough understanding of needs of the healthcare organizations. Expertise received with using strategies of records mining may be used to make successful selections in order to improve fulfillment of healthcare organization and health of the patients. Statistics mining calls for appropriate era and analytical strategies, aswell as systems for reporting and monitoring that may enable measuring of outcomes. On this paintings, Rule based totally mining gadget is used to expect the risk factor for cardiac arrest based totally on the consumer's primary profile and fundamental clinical details. Every element is assigned with weightage value. Risk thing prediction is based totally on the average weight price. This look at to be extended and tested with patient's scientific test info, social and economic information are taken to are expecting the diabetic and cardiac related issues.

REFERENCES

1. Fayyad U, Piatetsky-Shapiro G, Smyth P. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*.1996; 17(3):37-54.
2. Jenzi IS, Priyanka P, Alli P. A Reliable Classifier Model Using Data Mining Approach for Heart Disease Prediction. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2013 Mar; 3(3):20-4.
3. Shaorong F, Zhixue H. An Incremental Associative Classification Algorithm used for Malware Detection. *Proceedings of International Conference on Future Computer and Communication*. 2010; 1:757-60. crossref.
4. Cheng CW, Chanani N, Venugoplan J, Maher K. icuARM-An ICU Clinical Decision Support System Using Association Rule Mining. *IEEE Journal of Translational Engineering in Health and Medicine*. 2013; 1:1-10. crossref.
5. Leong KI, Si YW, Biuk-Aghai RP, Fong S. Contact tracing in healthcare digital ecosystems for infectious disease control and quarantine management. *Proceedings of International Conference on Digital Ecosystems and Technologies*. 2009; p. 306-11.
6. Balasubramanian T, Umarani R. An analysis on the impact of fluoride in human health (dental) using clustering data mining technique. *Proceedings of Pattern Recognition, Informatics and Medical Engineering (PRIME)*. 2012; p. 370-5.
7. Singh K, Upadhyaya S. Outlier detection: applications and techniques. *International Journal of Computer Science Issues*. 2012 Jan; 9(1):307-23.
8. Rahaman B, Shashi M. Sequential mining equips e-Health with Knowledge for mining diabetes. *New Trends in Information Science and Service Science*. 2010; p. 65-71.
9. Chignell M, Rouzbahman M, Kealey R, Samavi R, Yu E, Sieminowski T. Nonconfidential patient types in emergency clinical decision support. *IEEE Security & Privacy*. 2013; 11(6):12-8. crossref.
10. Cios KJ, William Moore G. Uniqueness of medical data mining. *Journal of Artificial Intelligence in Medicine*. 2002 Sep-Oct; 26(1-2):1-24. crossref.

