Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

EARLY PREDICTION AND DIAGNOSES OF CARDIO DISEASES DUE TO DIABETES MALLITUS: CARDIODIBET HYBRID MODEL

MALTI NAGLE
Department of Computer Science & Engineering, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India. Email: Nagle.malti083@gmail.com

PRAKASH KUMAR
Associate Professor, Department of Computer Science & Engineering, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India. Email: Prakash.kumar@jiit.ac.in

Abstract

A remarkable innovation in healthcare system presided to production of data of diabetes mellitus and cardio vascular diseases. Majority of population over the globe is suffering from one or the other kind of health issues. To visualize medical problems globally and to connect medical centre and patient all together, information technology contributes to excel in modern healthcare services. One, such advance IT field is Internet of things (IoT). IoT is step ahead in all aspects and so in healthcare. Health sort of area is something wherein IoT plays an important role. Inspite of that, IoT is not in mainstream adoption in healthcare sector, therefore it is required to develop enhanced version of traditional healthcare system that can facilitate to process data received from IoT healthcare devices. In this paper, (**cardiac+diabetes)** healthcare framework has been proposed to diagnose effects of diabetes over cardiac patients. Patient's data is investigated on the basis of findings of (diabetes + cardiac) clinical diagnosis profile. Novel framework consists of two methodologies, first is **GEETN** technique (pre-processing method) and the other is **Advanced Hoeffding** to train and test model**)**. Novelty of work is to pre-process (**cardiac+diabetes)** dataset using **GEETN**. The accumulated comparison is based on outcomes that consist of various algorithm with **Advanced Hoeffding lincasingan accuracy of 98.99%** Precision (92.99%), recall (97.56 %) and f1 score (95%) that helps in early detection of patients' health condition to reduce the rate of death cases, **cardiodibet** healthcare systems helps in providing better monitoring, communication and early diagnosis of diabetes and cardiac health of patients. The proposed method identify the preliminary status of diabetes and cardiac vascular diseases parameters of patient through Normal, Moderate and high risk further message is generated for critical or emergency cases. It also helps to identify the possibilities of silent heart attack of patients at early stage, consequently reducing the can reduce the number of death cases.

Keywords: Cardiodibet, Adaptive Hoeffding Algorithm, GEETN, Diabetes, Cardio Vascular Diseases. Hb1Ac, Hypertension, Hyperglycaemia.

1. INTRODUCTION

In today's scenario, majority of the diseases are on the rise due to sustained stress. Hectic lifestyle is also adding stress in everyone's life resulting in fatal health problems. Stress related issues are giving birth to severe health diseases like Diabetes, hypertension and weaken heart. Now a days, cardio vascular diseases are widely noted in majority of populations [1]. It has also become prevalent in early age of populations. Doctors and practitioners conduct inclusive survey and consolidate the data for heart related diagnosis and study. It has been observed from that cardio vascular diseases in patients are increasing at large. Many doctors state that traditional model is taking long time to diagnose. Sometimes results of patients' cardiac test comes after the cardiac arrest. All

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

of these are major challenges in saving patients' lives. The data gathered from survey of cardiac patients consist noticeable facts about their health status. If the symptoms can be diagnosed at early stages by any prediction model, it will be highly beneficial for patients to take timely consultation and cure for the heart diseases (like cardio vascular arrest, Heart coronary diseases, etc.).

In modern healthcare systems, various sensors are installed in patients' body, that can continuously read data from sensors and keep forwarding this data at cloud to process and analyse.[6] This real time health data is difficult to extract such massive and large dataset. Sometimes these information are in the hidden format. Streaming of such massive data is vital challenge for researchers. Techniques applied on real time data streaming can facilitate researchers attain two insights: novel approach and solicitous extraordinary perception of enormous health related datasets. Prediction and labelling of the diseases received through real-time data communication method is a tedious task.[7,8] This paper, applied various high precision and high accuracy algorithm to the cardiac related dataset to acquire the unidentified trends observed in heart patients. A dataset of 997 patients consisting of 8 attributes are taken into consideration to examine and find the accuracy for different algorithms. Accuracy and recall kappa values are examined and health status of patient is categorised based on dataset value. Further, Supervised learning technique is applied to refine the dataset and then accuracy is measured. The complete dataset is further categorised into three stages (normal, borderline and high risk) of cardiac strokes.

2. BACKGROUND STUDY

Stress related health diseases are very prevalent due to improper routine and busy schedule of every human being. It is escalating the chances of critical medical issues in person. Some of them are incurable diseases for example cardiac stroke and diabetes are very commonly observed in patients. This motivated the researchers to understand the hidden pattern of medical datasets. This section enlightens the study of real time medical dataset to acknowledge the heart diseases on the basis of their medical diagnosis. Data mining techniques that are used in previous works are also addressed in this section. [8, 9] The gap identified in existing works motivated to introduce suitable techniques that can classify the dataset to address the issues related to heart diseases. World organization stated that cardiac diseases are the most significant challenge majority of population are facing in recent days [15, 16].

In different paper several machine learning methods applied on various dataset to check false positive and true positives [18]. In literature survey, it has been observed that, the collection of data is part of cardiovascular disease retrospective studies utilizing the recordings of multichannel MCG. Many people with coronary stenosis and people who are healthy in the database. There are 16 NSTEMI (non-ST-elevation myocardial infarction) instances in the sample. For the ischemic group, coronary angiography is performed [19]. The coronary heart diseases can be predicted using ensemble model which can be helpful for early prediction of critical heart risks [23].

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

- Theerthagiri et al. presented machine learning algorithm to classify the diabetes using different supervised learning model. The multilayer perceptron along with decision tree and other algorithm has been used in the work. The prediction shown in research has defined the lowest false positive and lowest false negative rates under curve of 86% [1].
- Nagavelli et al. propsed model to detect heart diseases using ECG. Specific signal range of ECG has been captured for observation 20 number of times.[20] Readings of ECG has been encapsulated based on different probabilities of ECG signals reach to Peak and low waves of patients' ECG [17, 21, 22].
- Butt. et al. developed model to predict and classify diabetes of PIMA patients. Proposed methodology significantly classify diabetes patient with 86% accuracy and predict diabetes with 87% accuracy.[2]
- Das et al. presented classifier algorithm (Neural Network) as a solution for finding is an ensemble-based model proposed as a new model which deals with cardio vascular related issues, combination of predicted value and probability of multiple predecessors. The successful implementation and testing was obscured for 215 patients' dataset and an accuracy of 97.4% was achieved [3].
- Mirmozaffari et al. introduced model that uses different classifier on 209 instances of heart related dataset. All algorithm were tested on WEKA tool and achieved 97.6077% of accuracy by random tree classification algorithm. Unsupervised and supervised learnings were used to filter the data [4].
- Jia explained the VFDT algorithm to classify real time data stream on various type of datasets. This paper introduced the optimized technique to reduce the concept of drift. The result shown in paper clarify the effective and improved algorithm with less error in stream mining [5].
- Thaiparnit et al. proposed model "Vertical Hoeffding Decision tree. Sample of 190 patients were taken for analysis. Data mining applied for data preparation and selection. The result showed the accuracy of 85.43% and error rate of 14.5%. The smallest root expected at 0.3666, with 10 cross fold" [6].

3. DATASET

The dataset considered for the analysis comprises of diabetes and cardiac related records based on the cholesterol HDL and LDL long with insulin values. Data is collected from association of peripheral adipose tissues pathology with type 2 diabetes in Asian Indians. It consist of 60 records and "without any alteration [12]. It has total 9 attributes (8 input, 1 output) related to prescribe heart and diabetes diagnosis based on blood sample taken from patients. Table 1 consists of the attributes and its specifications.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

Table 1: Attributes of health patient

Attributes	Specifications
Gender	The gender specification of patient
Phenotype	The normal glucose tolerance test a genotype of patient based on blood sample
Triglycerides	It is a type of fat. It gets reserved in body from calories that are not burned in routine lifecycle and are stored in body as extra calories.
Cholesterol	It is a fat like waxy substance that gets accumulated in all cells and required to generate healthy cells and hormones in blood. But if it is in excess then this Attributes creates problem in blood flow rate and gives rise to the heart diseases. Sometimes it form clot in blood vain that leads to heart stroke
HDL	Stand for high-density lipoproteins. It is type of good cholesterol because it does not remains in blood for long time. HDL sends by body parts to liver and liver eliminate it from body.
LDL	It known as Low- density lipoproteins. It is also called 'bad' cholesterol it gets accumulated around the arteries.
VLDL	it stands for Very low density lipoproteins. It contains high amount of triglycerides and is also considered as one of the reason to buildup cholesterol around the arteries walls.
Insulin	Insulin is a form of hormone that let the body to use glucose available in food to produce energy. It is situated between pancreas and gland that is at the back side of stomach.it balance the glucose level and does not emits from the body it is stored in liver as it is.
Outcome	This attribute used to check the health status of one's heart based on the attributes collected from blood strains.

4. EXPERIMENTAL SETUP

Proposed model diobocardio consist of two phases: 1. **GEETN (**Pre-processing technique),

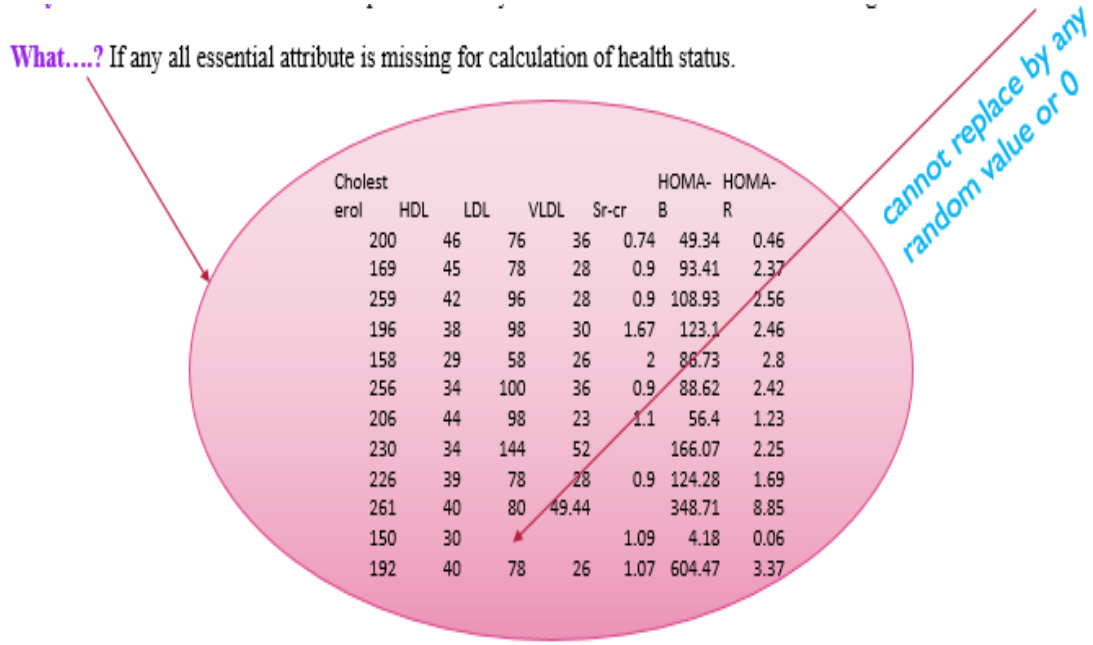
2. **Adaptive Hoeffding Algorithm** (Train and test Model). To accomplish the diagnosis of Diabetes + cardiac profile of patients both the model contributed successfully to produce efficient outcomes.

4.1. Pre- processing technique (GEETN) –

This proposed novel technique GEETN) is to pre-process the raw data consisting of findings of cardio vascular and diabetes. At first all the missing values were identified in raw healthcare data. Later it has been replaced with the standard values of attributes based on the standard healthcare defined by medical research center.

Pre – processing – algorithm (GEETN) - The missing and null value of dataset has been changed on the basis of standard value of attributes defined by doctor. Second challenge is to process real time health data streaming. Handle the null value is also matter of concern. Because any value in health data is specific health reading Why…..? – Health attributes can be replaced with any random value and not with mean average value. It can't be 0 either.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786



Cholesterol	HDL	LDL	VLDL	Sr-cr	HOMA-B	HOMA-R
200	46	76	36	0.74	49.34	0.46
169	45	78	28	0.9	93.41	2.37
259	42	96	28	0.9	108.93	2.56
196	38	98	30	1.67	123.1	2.46
158	29	58	26	2	86.73	2.8
256	34	100	36	0.9	88.62	2.42
206	44	98	23	1.1	56.4	1.23
230	34	144	52		166.07	2.25
226	39	78	28	0.9	124.28	1.69
261	40	80	49.44		348.71	8.85
150	30			1.09	4.18	0.06
192	40	78	26	1.07	604.47	3.37

Figure 2: Categorical Dataset

The following steps is applied to calculate GEETN.

1) All value of attributes has been calculated for its normal range of health parameters.
2) Health dataset is pre-processed using the normal range of each attributes. Then the outcomes is calculated according to the value of each attribute.
3) The proposed method then identify as the null value in healthcare dataset and the healthcare dataset based on the novel approach.
4) Imputation method for categorical columns are applied.
5) Impute the missing value using GEETN.
6) The missing value is first detected and replaced by calculating values based on other parameters.
7) GEETN is applied on attributes those are correlated with other attributes. For example cholesterol value predicted by blood strain is correlated with (LDL,HDL,VLDL). Likewise Hb1ac.
8) Imputation is evaluated on the basis of health value of different attributes.
9) Dependent variable is calculated by using following equation:

$$(output)\text{^}n = \sum\_(i = 0)\text{^}n \text{▒} \llbracket (\text{input\_attributes}) \rrbracket$$

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

10) Further the output (dependent variable) is labelled according to the range of abnormality in diagnosis. Equation to label dependent variable:

TN= Total number of labels considered in range of abnormality

y= average of TN

z= $\sum_{y}^{TN} z$

$$|label_{outcome}: x| = \begin{cases} x = 0, & x = 0 \\ x = 1, & x > 0, x \leq y \\ x = 2, & x \geq z, x \leq z \end{cases}$$

11) Output is the new leaf node attribute from all observations of input_attributes that has been defined for normal health of patients.

$(output)^n = \sum_{i=0}^{n} (\text{input\_attributes})$ For normal healthy readings of patient.

4.2 Proposed Algorithm

Adaptive Hoeffding Algorithm

This paper proposed the prediction model for real time stream data prediction. The real time Adaptive Hoeffding tree works on concept of incremental, the algorithm is useful to learn from real time streamed massive data. Examples generated from distribution of data streams are static with time. Therefore, Hoeffding Tree conclude the theory that, optimal split can be achieved by sometimes small sample from whole dataset. The Hoeffding bound explains this concept in algorithm, which state that estimation of given precision up to some extent require statistics of observed values [6].

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

Flowchart

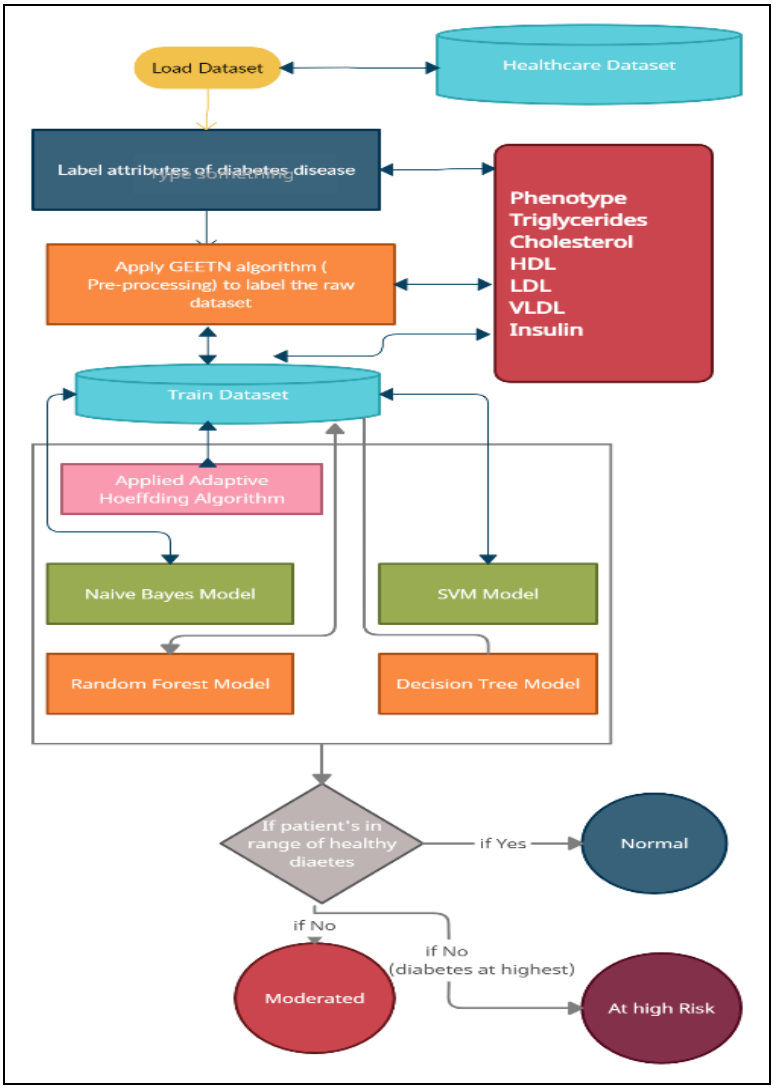


Figure 1: Flowchart of Adaptive Hoeffding Algorithm

Figure 1 shows the flow chart of the algorithm where the process consist various steps. The process of filter the data is based on statistics and list is sorted towards the leaf node splitting of dataset is completed by arguments values.

Hoeffding trees supports evaluation of rapidly growing real time streamed data. Proposed algorithm succeed to necessitate such data streams in comparison to traditional one, upside outcomes was exceptional to its introductions [7].

Hoeffding tree algorithm gives extraordinary results from traditional decision tree by elevating data stream exponentially. It inspect each sample of data stream just one time, the sample which employed in tree generation will be exempted from list of examples. Final tree will be stored memory of algorithm. It acknowledge only necessary samples

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

which participate in tree generation, leaves that helps in grow of tree. This sample further facilitate prediction at certain point of time to form training model [8].

Algorithm Adaptive Hoeffding Tree AHT classifier

```
1: AHT first node (start node)
2: repeat till al best value chosen for training data
3:     compare with existing node and replace (if greatest value found from dataset)
and make it as next leaf node of AHT.
4:     Update iteration value of incremental variable (i)
5:     increase count of num(i) for given sample in dataset
6:     If modulus[num(i), minvalue(num) - 0 & dataset[I] does not match with all
class of data then
7:     Calculate Gini_value(Xhti) for all elements from dataset
8:     choose Xhta with greatest value from elements
9:     consider Xhtb element next greatest value Gini_value
10:    evaluate AHT bound_value using equation 2
11:    if Xhta ? Xhtb and (Gini_value (Xhta) - (Gini_value (Xhtb) > AHT_bound or
AHT_bound < τ) then
12:       check split level of Xhta and select node from new set to Replace with I
13:       repeat split step for all possible branches
14:          if new node found meet expected outcomes then add as leaf node for new
level
15:       end step 13
16:    end step 11
17:      end step 6
18:   end step 2
```

Adaptive Hoeffding Tree Classifier

1) Gini_value has been calculated by taking difference of average value of wght_en , before split process start the value of entropy of distributed class.
2) The AHT_bound conclude difference between R(true mean) and observed mean for n should be less than: $\epsilon = \sqrt{\frac{R^2 \ln(\frac{1}{\delta})}{2n}}$ ……Eq -2
3) Gini_value for R range is log base (2) for all possible classs lebels.
4) Other parameters has been calculated using standard Hoeffding algorithm. Tie-breaking situations have been managed by checking the AHT_bound value and try to separate them.

5. RESULT AND DISCUSSION

Proposed algorithm introduced advance healthcare system that can successfully predict diabetes patient's health status. It also helps to predict heart health and cardiac arrest chances at early stage which is very critical to predict in traditional healthcare system. It was observed by doctors that due high diabetes patient died due to silent cardiac arrest. This cardodibet can be useful to predict early stage of cardiac arrest.

The diabetes detection health system generated using prediction of real time health data stream consist of diabetes and heart related attributes as it is combinational dataset, where attributes are dependent and missing value calculation is key challenge. Therefore,

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

it requires pre-processing algorithm by applying imputation methods. The proposed work contributed GEETN pre-processing algorithm to prepare training dataset after pre-processing

The pre-processed data were further used for prediction. Comparison outcome were illustrated by comparing various prediction algorithms Naïve bayes Decision Tree, Random Forest SVM with proposed Adaptive Hoeffding algorithm on the cardiac related dataset [9,10,12]. The pre-processing of is conducted by comparing the data for the healthy range of cardiac blood sample's readings. Outcomes of cardiodibet suggested that proposed classifier gives better performance to filter the dataset and provides the health status of cardiac patient with more accuracy. Handling real time data streaming of the missing value for health data is also the challenge that could overcome by GEETN method. Table 2 shown description of Adaptive Hoeffding algorithm with different parameters - accuracy of proposed method is 98.99%. Precision, recall and f1 score (92.99%, 97.56 and 95% respectively) is also higher than other traditional algorithm. Recall value is 97.56%. Overall performance of Adaptive Hoeffding is outperform as compared to other traditional algorithm. The graphical representation of all parameters shown in figure 4 is well defined. Along with performance the proposed method also categorise the number of patients according to their health status Normal, Moderate and High risk of their health status on the basis of prediction. Figure 3 shows the counts the health status of patients by checking the risk factor evaluated by attributes of diabetes and cardiac symptoms reading.

Hoeffding algorithm proposed in developing cardiodibet supports to provide outcomes of real time healthcare data, Following parameters have been produced to analyse the outcomes of diabetes + cardiac data.

Matrix consists of – accuracy, recall, kappa, f1 score are calculated using label_outcome as output attribute and patient's medical readings input attributes. Figure 4 represents the visualization of performance of cardiodibet healthcare system

Table 2: Overall Performance of Various Algorithm

Confusion Matrix	Naïve Baye	Decision Tree	Random Forest	SVM	ADAPTIVE Hoffding
Accuracy	84.75	94.99	90.88	88.48	98.99
Precision	68.18	70.24	73.56	77.88	92.99
Recall	89.47	91.26	91.36	94.88	97.56
f1	88.31	92.88	93.76	95.88	95.6

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

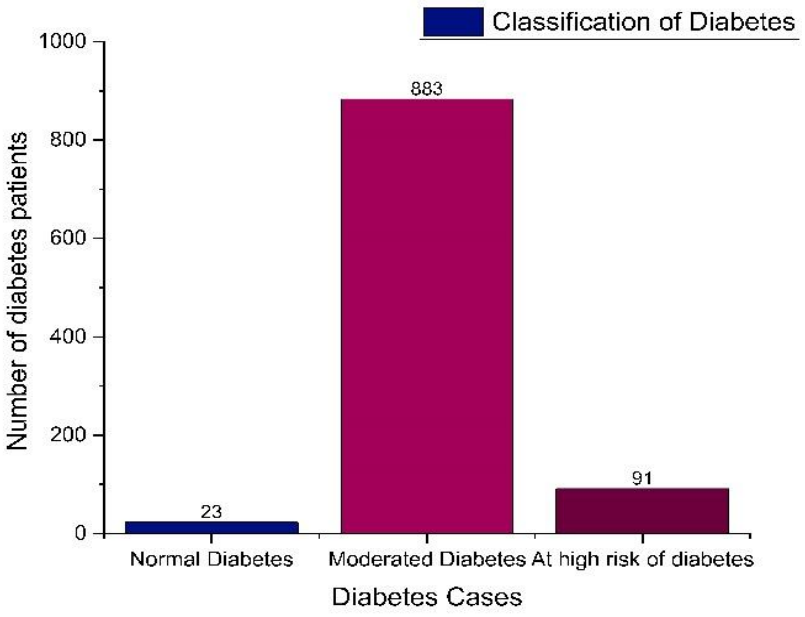
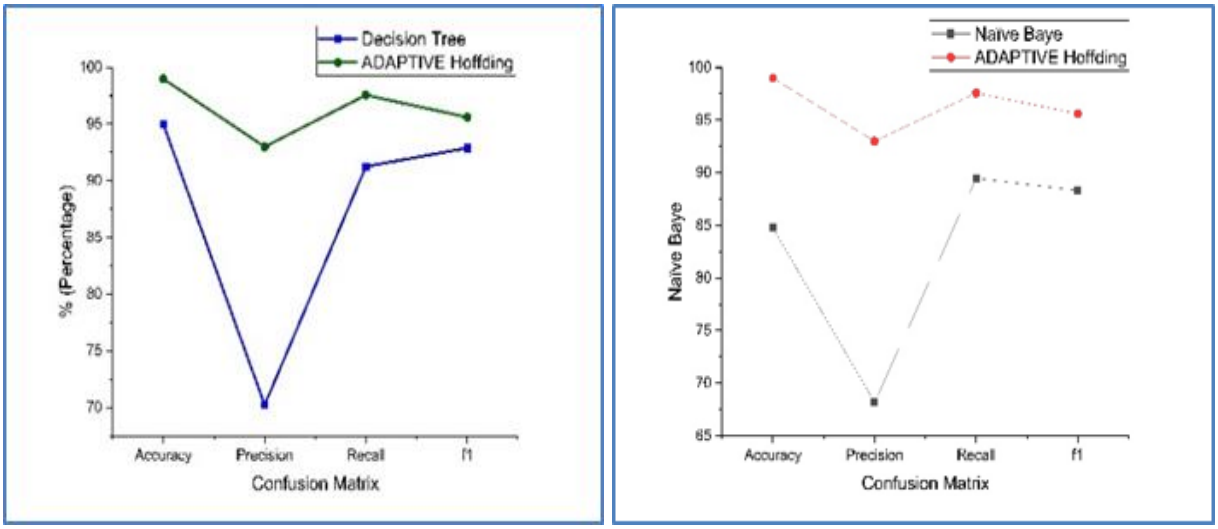


Figure 3: The results categorised the health status of patients



Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

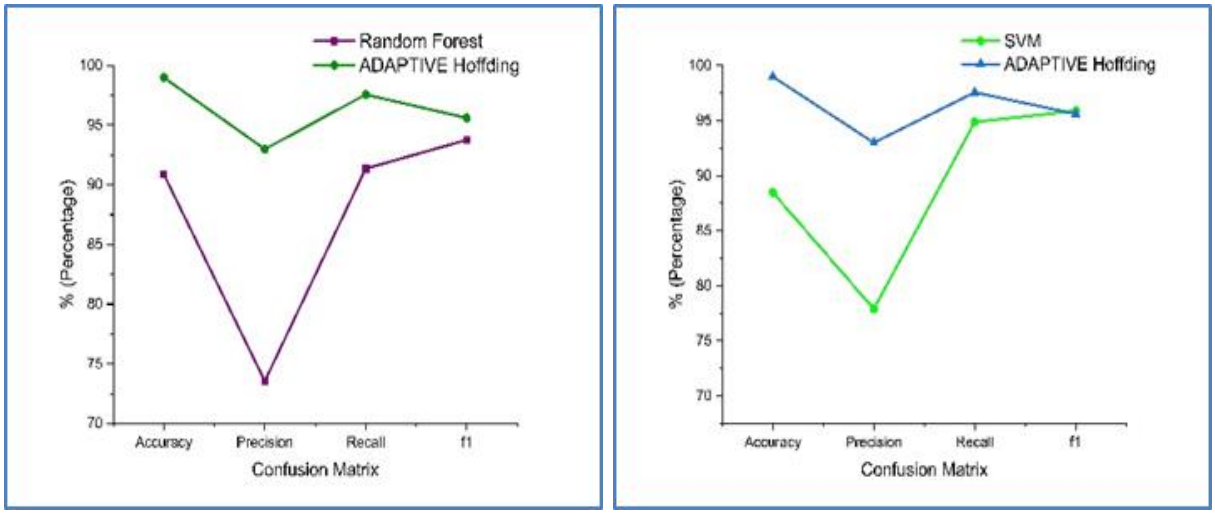


Figure 4: Comparison results of confusion matrix for proposed work

6. CONCLUSION AND FUTURE SCOPE

To detect the healthcare real time data (Based on Diabetes and cardiac diseases), algorithm applied on dataset those are best for feature prediction and splitting the dataset to predict the disease accurately. Comparison of traditional algortihm and Adaptive Hoeffding applied for prediction of health disease. Adaptive Hoeffding algorithm with different parameters - accuracy of proposed method is 98.99%. Precision, recall and f1 score (92.99%, 97.56 and 95% respectively) is also higher than other traditional algorithm. Recall value is 97.56%. Overall performance of Adaptive Hoeffding is outshine than as compared to traditional algorithms. The graphical representation of all parameters shown in figure 4 is well defined. Along with the performance, the proposed method also categorises the number of patients according to their health status (Normal, Moderate and High risk) using proposed methodology. Novel approach applied in proposed work counts the health status of patients by checking the risk factor evaluated by attributes of diabetes and cardiac symptoms reading. Adaptive Hoeffding algorithm gives 98.99% accuracy helps to predict the health disease at early stage based on patient's data and precautionary measures may be suggested by doctors. The cardiodibet is based on prediction of cardiac + diabetes symptoms the classification of patient's health status can be improved.

Highlights

1. The article is prediction of cardiac arrest based on impact of high diabetes.
2. Article incorporated related work to show the state-of-the-art of best suitable prediction supervised learning models for prediction.
3. Diagnosis of Blood Glucose level, LDL,VLDL and HDL are helpful to find impact of high sugar on cardiac vascular arteries.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

4. The implementation is carried out for the prediction of cardiac vascular patients using Hoeffding Algorithm and GEETN proposed methods for training the dataset.
5. The prediction model is capable to categories the health status of patients into three catogires (Normal diabetes patient, Moderate, High risk of cardiac arrest).

References

1) P. Theerthagiri, A Ruby, L. Vidya, “Diagnosis and classification of the Diabetes Uaing Machine Learning Algorithms”, SN Computer Science, Springer Nature, 2022,
2) U.M. Butt, S. Letchmunan, M. Ali, F. H. Hassan, A. Baqir, H.H.R Sherazi, “Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications”, Journal of Healthcare Engineering, Oct-2022, Vol. – 2021, PP 1-17. https://doi.org/10.1155/2021/9930985
3) R. Das, I. Turkoglu, and A. Sengur, “Diagnosis of valvular heart disease through neural networks ensembles,” Elsevier, 2009.
4) M. Mirmozaffari, A. Alinezhad, A. Gilanpour, “Data Mining Classification algorithm for Heart disease prediction”, (IJCCIE), 2017 vol 4, pp. 11-15.
5) S. Jia,“A VFDT algorithm optimization and application thereof in datastream classification”, ICAMLDS,2020.
6) S. thaiparnit, S. Kritsanasung, N. Chumuang, “A Classification for Patients with Heart Disease Based on Hoeffding Tree”, 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2019, p.no. 2372-164.
7) P. Domingos and G. Hulten. Mining High-Speed Data Streams. In KDD, pages 71-80, Boston, MA, 2000. ACM Press.
8) G. Hulten, L. Spencer, and P. Domingos. Mining time-changing data streams. In KDD, pages 97–106, San Francisco, CA, 2001. ACM Press.
9) Bifet, A., Holmes, G., Kirkby, R. and Pfahringer, B., 2010. Moa: Massive online analysis. Journal of Machine Learning Research, 2011(May), pp.1601-1604.
10) Bifet, Albert, and Ricard Gavaldà. "Adaptive learning from evolving data streams." International Symposium on Intelligent Data Analysis. Springer, Berlin, Heidelberg, 2009.
11) Sharma, Himani, and Sunil Kumar. "A survey on decision tree algorithms of classification in data mining." International Journal of Science and Research (IJSR) 5.4 (2016): 2094-2097.
12) Rish, Irina. "An empirical study of the naive Bayes classifier." IJCAI 2001 workshop on empirical methods in artificial intelligence. Vol. 3. No. 22. 2001.
13) Hoeglinger, Stefan, and Russel Pears. "Use of hoeffding trees in concept based data stream mining." 2007 Third International Conference on Information and Automation for Sustainability. IEEE, 2007.
14) Kumar, Arvind, Parminder Kaur, and Pratibha Sharma. "A Survey on Hoeffding Tree Stream Data Classification Algorithms." CPUH-Research Journal 1.2 (2015): 28-32.
15) Wang J., Liu C., Li L., et al. A stacking-based model for non-invasive detection of coronary heart disease. *IEEE Access*. 2020;8 doi: 10.1109/access.2020.2975377.37124 [CrossRef] [Google Scholar]
16) Zhenya Q., Zhang Z. A hybrid cost-sensitive ensemble for heart disease prediction. *BMC Medical Informatics and Decision Making*. 2021;21(1):p. 73. doi: 10.1186/s12911-021-01436-7. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 56 Issue: 12:2023
DOI: 10.5281/zenodo.10426786

17) Nourmohammadi-Khiarak J., Feizi-Derakhshi M.-R., Behrouzi K., Mazaheri S., Zamani-Harghalani Y., Tayebi R. M. New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. *Health Technology*. 2020;10(3):667–678. doi: 10.1007/s12553-019-00396-3. [CrossRef] [Google Scholar].

18) Gomathy V., Padhy N., Samanta D., Sivaram M., Jain V., Amiri I. S. Malicious node detection using heterogeneous cluster based secure routing protocol (HCBS) in wireless adhoc sensor networks. *Journal of Ambient Intelligence and Humanized Computing*. 2020;11(11):4995–5001. doi: 10.1007/s12652-020-01797-3. [CrossRef] [Google Scholar]

19) Thanga Selvi R., Muthulakshmi I. An optimal artificial neural network based big data application for heart disease diagnosis and classification model. *Journal of Ambient Intelligence and Humanized Computing*. 2021;12(6):6129–6139. doi: 10.1007/s12652-020-02181-x. [CrossRef] [Google Scholar]

20) Biswal A. K., Singh D., Pattanayak B. K., Samanta D., Chaudhry S. A., Irshad A. Adaptive fault-tolerant system and optimal power allocation for smart vehicles in smart cities using controller area network. *Security and Communication Networks*. 2021;2021 doi: 10.1155/2021/2147958.e2147958 [CrossRef] [Google Scholar]

21) Sultan Bin Habib A.-Z., Tasnim T., Billah M. M. A study on coronary disease prediction using boosting-based ensemble machine learning approaches. Proceedings of the 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET); December 2019; Dhaka, Bangladesh. pp. 1–6. [CrossRef] [Google Scholar]

22) Shorewala V. Early detection of coronary heart disease using ensemble techniques. *Informatics in Medicine Unlocked*. 2021;26 doi: 10.1016/j.imu.2021.100655.100655 [CrossRef] [Google Scholar]

23) Guo A., Pasque M., Loh F., Mann D. L., Payne P. R. O. Heart failure diagnosis, readmission, and mortality prediction using machine learning and artificial intelligence models. *Current Epidemiology Reports*. 2020;7(4):212–219. doi: 10.1007/s40471-020-00259-w. [CrossRef] [Google Scholar]