Body Health Monitoring System using Advanced Data Mining

R. Elakya [1], Naresh Kumar [2], Anand Anil [3], R. Rajasekar [4], Santhosh [5]
[1] Asst. Professor, CSE Department, SRM University, Chennai, Tamil Nadu, India.
[2, 3, 4, 5] UG Scholar, CSE Department, SRM University, Chennai, Tamil Nadu, India.

Abstract – We propose a advance Health Monitoring System used specially to monitor and to warn a patient in advance regarding anomalies in his health condition, with the help of PPG and PPT sensors and a user end device supported with advanced data mining techniques like Multi Task Learning, Preprocessing Scheme for High Cardinality Categorised Attributes to classify and predict problems based on an Deep or Multilevel Artificial Neural Network .

Index Terms – Artificial Neural Network (ANN), PPG Sensor, PPT Sensor Multi Task Learning(MTL),High Cardinality Categorised Attributes

1. INTRODUCTION

Todays fast lifestyle prevents people from taking good care of them which leads to a variety of health perils causing grave danger to the life of the people. Some of the major diseases are

- Hearth Diseases
- Blood Pressure and Sugar levels
- Respiratory diseases

A person could get them tested and have the results once he feels unwell but the results may take some time and there is no way to know the attacks or symptoms of the disease before-hand other than the gut feeling of a person .So we thought about a piece of tech which could be implanted in a human body which could analyse and and communicate the body parameters with an external agent which could be used to communicate with high end servers that have huge computing power to respond in realtime where the main advantage is its huge computational and comparable technique with previous datas of the same users in realtime.We would like to improve this with the help of advanced data mining techniques along with the help of Artificial Neural Network.To improve health care with a minute but critical step is the major duties of an professional ,which we try to fulfil here by proposing this advancement of an existing system.

2. RELATED WORK

There exist a similar system proposed by Johan Wannenburg to monitor and diagnose the health and vitals of a person when implanted with but the effectiveness regarding the data mining and which could not prove to be as worthy compared to when upgraded with advanced data mining techniques. The existing System doesn't have any major fault but is still a bit far from reality and economically possibility ,so we shave tried to improve the existing system to benefit the users more with an more accurate and personalised result and diagnosis. We have used the following ideas to bring together our proposed system

- Photo-plethysmography
- Electrocardiography
- Pulse Transit Time Detection Process
- A Preprocessing Scheme for High-Cardinality Categorical Attributes in Classification and Prediction Problems
- Multi Task Learning

2.1. Photo-plethysmography - the process to optically obtain plethysmogram, a volumetric measurement of an organ. A PPG is often obtained by using a pulse oximeter which illuminates the skin and measures changes in light absorption. A conventional pulse oximeter monitors the perfusion of blood to the dermis and subcutaneous tissue of the skin. PPG (photo-plethysmography) sensors use a light-based technology to sense the rate of blood flow as controlled by the heart's pumping action

2.2. Electrocardiography - is the process of recording the electrical activity of the heart over a period of time using electrodes placed on the skin. These electrodes detect the tiny electrical changes on the skin that arise from the heart muscle's electrophysiologic pattern of depolarising and depolarising during each heartbeat. ECG (electrocardiography) sensors that measure the bio-potential generated by electrical signals that control the expansion and contraction of heart chambers.

2.3. Pulse Transit Time Detection Process - The process of blood pressure extraction will also make use of the pulse oximetry hardware, once the PPG signals are digitised, they are processed to detect PTT. Other implementations make use of both PPG and ECG sensors to do this but here two PPG sensors are used. The two PPG sensors are set up in such a way that one is leading and the other lagging. Both sensors measure PPG signals, however due to their location and relevant distances from the heart the

one signal leads the other in a time domain measurement. This time is measured using designed software algorithms and used in the calculation and prediction of blood pressure.

2.4. A Preprocessing Scheme for High-Cardinality Categorical Attributes in Classification and Prediction Problems - The method leverages a robust estimation formula that finds its roots in the Empirical Bayesian framework. Because the method basically requires only aggregation functions, it can be easily carried out via native database operations. This represents a significant advantage compared to clustering methods, which require more complex algorithms and also lead to greater information reduction. Furthermore, the hierarchical version of the proposed method takes advantage of categorical data that can be aggregated to different levels of granularity, providing an effective data-driven variant aggregation scheme. A brief review of the statistical and machine learning literature has shown that, while shrinking the observed probabilities toward the prior probabilities is the common theme, different forms of weighting factor $\lambda(n)$ have been proposed. Thus, a comparative experimental analysis of the various forms of $\lambda(n)$ is warranted and it will be addressed in subsequent publications. Although probability estimation methods are nearly 50 years old, their adoption in the machine learning community is only relatively recent. Furthermore, specifically in the neural network community little has been done to address the use of high-cardinality categorical attributes with these types of models. Indeed, often practitioners simply ignore these data fields although they could provide significant value to their models. As this paper has hopefully shown, high-cardinality categorical attributes can be included in any predictive model via a statistically sound transformation that it relatively straightforward to implement. This is one more example of how well established statistical methods complement modern data mining techniques.

2.5. Multi Task Learning - MTL is a subfield of machine learning in which multiple learning tasks are solved at the same time, while exploiting commonalities and differences across tasks. This can result in improved learning efficiency and prediction accuracy for the task-specific models, when compared to training the models separately. MTL aims to improve the performance of multiple classification tasks by learning them jointly. One example is a spam-filter, which can be treated as distinct but related classification tasks across different users.

3. PORPOSED MODELLING

We here emphasise and propose the similar hardware and configuration as proposed by Johan Wannenburg ,the model could certainly improve when also concentrated upon the data mining and analysis part which could increase the efficiency of the system to accurately and personalise the results according to the user's body . The architecture of the system is provided below:

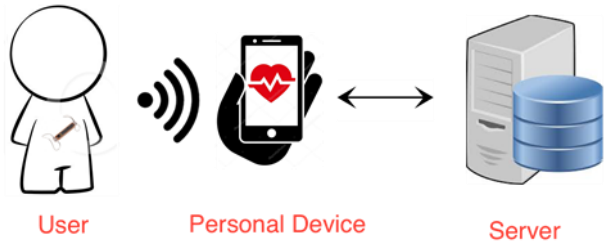


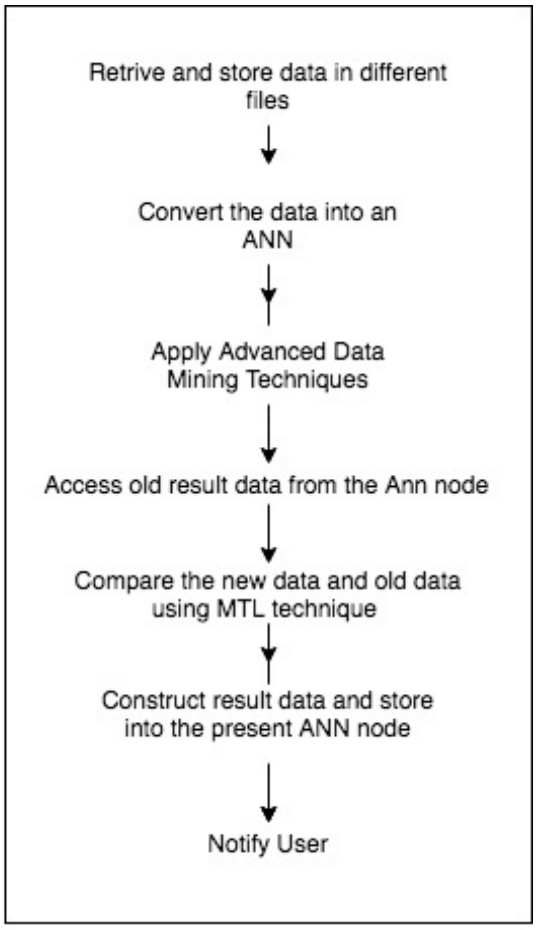
Figure 1 System Architecture

The sensor collects the details like Heart beat rate along with the necessary vitals of the body monitored and sends it to the nearby user personal device which has an access to internet and with credentials of the user so as to upload and retrieve results from the server .The basic tasks which are performed when data is taken to the server is shown below in the figure 2.

Here as we can see the data is stored in files which are used to map and create the ANN which is the vital memory part and is basically the backbone of the whole data mining process , without whom these processes would not be able to be performed . The data taken in as input has to be between an standard expected charts of data which are expected in any scenario in a human being unless he is dead , we could not have abnormally high heart rate or low heart rate and have the person still functioning which is why we can say the data attributes are of high cardinality but still needs be within limits to perform these functions .So when acquiring data these conditions are preset and an ANN with all the connection to the present node having time and date as reference frame and other health parameters as the attributes to the node ,this is how the ANN is constructed and maintained with the processing creating a resultant node or attribute connected to that present attribute when compared with the old data by creating an average parameters from all the previous data which is created with the help of MTL algorithms performed simultaneously alone with the preprocessing scheme on high cardinality categorised attributes . The preprocessing helps with the classification and prediction of the result but MTL creates the average median data from previous data .So we have to make sure this both algorithms run smoothly with interfering with each other making it a bit cumbersome task to solve but when achieved it will create marvellous results in data analysis and mining job.

The ANN helps with the high level association and mapping which could not be achieved by using Bayesian network as it fails on huge data mapping task and can't improvise as needed . Finally when the data is converted and stored in as an ANN ,we compare the data with old data using different graphical and range comparison methods which prove to be an edge

giving method to improve personal results as the median and average data of user is compared rather than the normal average



.

Figure 2 Server Tasks Undertaken

4. RESULTS AND DISCUSSIONS

The results could not be compared due to the limits of research but we are confident that the configuration proposed is of the best thinkable solution available to this data section. Comparatively less errors and higher accuracy as the data is compared and takes in accordance to the user and not the assumed median for general population .We have proposed the model with an upgraded data mining techniques which improve it drastically and give an better accuracy and efficiency along with personalisation.We could also argue that there could be other techniques that could be applied to improve the outputs and other factors but the techniques could not be applied to these high cardinality categorised attributes data .So hence we argue the method selected will be the best to map and do multi task learning.

5. CONCLUSION

The system depicts the results expected and could be put together to create an advanced system as an health monitor to predict and analyse Human body vitals. The different algorithms and schemas used along with ANN provide a better system to the users with high accuracy and could be further improved by researchers taking into consideration the different existing algorithms .We used these algorithms as they suit the data and don't neutralise each others effect on the data. This system could be future improved with better algorithms but still at present the best the algorithms could be combined.

REFERENCES

[1] Johan Wannenburg and Reza Malekian "Body Sensor Network for Mobile Health Monitoring, a Diagnosis and Anticipating System "
[2] Priyanka N, "Usage of Data mining techniques in predicting the Heart diseases – Naïve Bayes & Decision tree ", 2017 International Conference on circuits Power and Computing Technologies [ICCPCT].
[3] Mujtaba Ashraf Qureshi , Dr. Irshad Ahmad Mir , "COMPARATIVE STUDY OF EXISTING TECHNIQUES FOR HEART DISEASES PREDICTION USING DATA MINING APPROACH ,"Asian Journal of Computer Science And Information Technology 7: 3 July (2017) .
[4] Sarath Babu, Vivek EM, Famina KP, Fida K, Aswathi P, Shania M, Hena M , "Heart Disease Diagnosis Using Data Mining Technique ," International Conference on Electronics, Communication and Aerospace Technology ICECA 2017
[5] Salma Banu N.K, Suma Swamy, "Prediction of Heart Disease at early stage using Data Mining and Big Data Analytics ," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT) .
[6] Daniele Micci-Barreca "A Preprocessing Scheme for High-Cardinality Categorical Attributes in Classification and Prediction Problems ," SIGKDD Explorations,Volume 3, Issue 1

Authors

Ms. R. Elakya a Assistant Professor (O.G) ,CSE Department at SRM university ,has an M.E. in Computer Science and Engineering from Anna University 2013 batch .Area of Interest Computer Networks

Naresh Kumar an UG scholar at SRM University ,CSE Department 2014-2018 batch with area of interest as Artificial Intelligence and Data Mining .Presented an paper "AN INTELLIGENT PERSONAL ASSISTANT
MODEL USING IMAGE PROCESSING AND
ADAPTIVE LEARNING" at NCSIRCS 2017

Anand Anil an UG scholar at SRM University ,CSE Department 2014-2018 batch

Santhosh an UG scholar at SRM University,CSE Department 2014-2018

Rajasekar an UG scholar at SRM University ,CSE Department 2014-2018 batch