Data Analytics on IoT-based Health Monitoring System

Mayank Murali, Mahima Bhargava, Snehaa G, Anuj Anand, Md Amaan Haque,

Vergin Raja Sarobin M

Abstract**:** ***In recent times, health problems like cardiac failure, lung failures and heart related diseases are arising day by day at a very high rate. To address these problems, a sustained health monitoring becomes imperative. By regular health monitoring, diseases can be diagnosed and treated early. A new concept which is emerging is by monitoring the health of a patient wirelessly. In this paper, we report the use of a body temperature sensor and a heartbeat sensor for an IoT-based health monitoring system. The set up involved relaying of data from the sensors to an Arduino which was then sent to the cloud via an Ethernet module. The collected data were further processed by applying various algorithms and classification was done. We present the results obtained from the processing of the data using various algorithms - Naive Bayes, Random Forest Tree and SVM- and their comparison based on accuracy, precision and classification time.***

Index Terms**:** ***Cloud Based Data Analytics, Health Monitoring, Internet of Things, Prediction and Classification.***

I. INTRODUCTION

The new trends in the lifestyle leading to unhealthy eating habits, etc. have caused the potential diseases particularly of the heart and vascular to increase dramatically. Moreover, it is seen that the younger generation has also started suffering from heart problems [1]. Most of the deaths caused across the globe are because of the coronary heart disease. Thus, any progress leading to improved early diagnosis in a patient is always welcomed by medical community.

Two most important parameters measured from a patient are the heartbeat and body temperature, out of which heartbeat is significant for the cardiovascular system. Heartbeat is described as the aggregate amount of heart beats per unit of time and is expressed in terms of beats per minute (bpm). On an average, a healthy adult has a heart rate of around 72 bpm [2]. Based on the intake of oxygen by the body, the heartbeat differs during exercise and at rest. People who are involved in

Training and exercise have lower heart rate compared to people who are less physically active. Heart rate races when a person exercises and it slowly returns back to normal rate after the physical exercise is completed. The rate at which the heart rate drops back to the normal value is an indicator of the fitness level of a person.

Pulse was common method of disease diagnosis by doctors in olden days. In modern times too, measured heart beats are utilized by medical experts as a fundamental basis to help in the diagnosis and tracking of the medical conditions [3]. It is also utilized by the individuals engaged in intense physical training, such as athletes, to achieve maximum efficiency, and can also be employed for monitoring the health of elderly people. Heart rate is usually measured in controlled environment in clinics, but a portable system which can be used to monitor health anywhere, anytime has become a necessity [4]. The system developed by us and reported in this communication, has the advantage of such portability.

Similarly, the body temperature of an individual is also important and has been found to vary throughout the day. The average temperature of an adult normal body is 37 0C or 98.6 0F [5]. During the early morning, the temperature is established to be the lowest and is the highest during the evening. These changes are important and needed to be measured. As the body temperature varied according to the time, it was difficult to accurately measure it. So, we used LM35 temperature sensor which generated an analog output voltage that is proportional to the temperature. Since this analog value could not be used in this form, this temperature sensor required an analog digital converter to convert the analog value to a digital form [6]. For this reason, we used Arduino Uno to convert the analog value to a digital form in order to send the measured data to the cloud.

In this paper, we present the design for an economical health monitoring system which will measure vital parameters such as body temperature and heart rate of an individual. This measured data is then transferred to cloud where it is stored. This stored data is then processed and analyzed for which we implemented different classification algorithms such as Regression Tree, Naive Bayes and SVM and compared the results of the same.

Revised Manuscript Received on May 18, 2019

Mayank Murali, B.Tech Student, School of Computing Science and Engineering,, Vellore Institute of Technology, Chennai, India.

Mahima Bhargava, B.Tech Student, School of Computing Science and Engineering,, Vellore Institute of Technology, Chennai, India.

Snehaa G, B.Tech Student, School of Computing Science and Engineering, Vellore Institute of Technology, Chennai, India.

Anuj Anand, B.Tech Student, School of Electrical and Electronics Engineering, Vellore Institute of Technology, Chennai, India.

Md Amaan Haque, B.Tech Student, School of Electrical and Electronics Engineering, Vellore Institute of Technology, Chennai, India.

Vergin Raja Sarobin M, Assistant Professor, School of Computing Science and Engineering, Vellore Institute of Technology, Chennai, India.

II. LITERATURE REVIEW

Parameters, namely, the body temperature heart rate and respiratory rate of a patient are sensed and this data is transferred to a raspberry pi. This data can be utilized for the diagnosis of patient anywhere in the world using internet [7]. This record of the various parameters can also be stored on cloud of medical history and is displayed in form of web page for easy access of data. Without the need of physical presence, doctors can interact with patient from anywhere and the system can provide some aid by automatically generating the graph of body changes.In another research, various vital parameters of a patient were measured. These included heart rate, plethysmography, heart rate variability, pulse oximetry, ECG and fall detection were measured and evaluated [8]. Two distinct design frameworks of a (Wireless) Body Area Network were proposed. Android phones can be connected with the body area network. It has real-time system features which can do data acquisition in the (W)BAN, analysis and data storage and visualization on the android device, along with emergency communication and data transmission with server. Smart energy efficient sensor nodes and a coordinator node were used for Zigbee based methodology. The sensor nodes were used for acquiring various physiological parameters, performing data analysis, signal processing and transmitting measurement values to the coordinator node. In another design, sensors were linked via cable to an embedded system. For both approaches, Bluetooth was utilized for transferring data to an Android-based Smartphone and several challenges like reliability of wireless data transmission, security and many more were also discussed.Survey was carried out on existing communication protocols and security problems related to widespread health monitoring, their challenges, limitations, and probable solutions were described [9]. This resulted in the proposal and development of a generic protocol stack design for managing interoperability in heterogeneous low-power wireless body area networks. The design of this system was based on Wireless Sensor Networks (WSN) and smart devices [10]. The doctor and the caregivers could observe the patient in real time, without requiring physical presence. The medical history and the health complications such as report and medications of the patient were uploaded to web server which could be retrieved at any point of time as per convenience. The cloud analyzed data provided more insight into the patient's health. SenML was implemented for the transmission of sensor metadata, which guaranteed organized behaviour. The privacy content of the history and medical data were also defined. Add-on services were also defined which included Action (ReTiHA), Real Time Health Advice and Parent monitoring. An application named 'ECG Android App' had been fabricated, which provided the end user with visualization of their Electro Cardiogram (ECG) waves along with a data logging functionality in the background [11]. This logged information could be uploaded to the user's private centralized cloud or a specific medical cloud, which was responsible for maintaining a record of all the monitored data and this data could be accessed by the medical staff for analysis. A raspberry pi microcomputer based system had been developed which gathered various sensor data such as ECG signal, blood pressure signal, heart beat signal, situation of oxygen in blood (SPO2), temperature and many other signals from the patient were sent to medical specialists who were physically away from the concerned patient [12]. An online application was developed to foster efficient communication between specialists and paramedics.

Verifiable comparisons were administered on various supervised algorithms such as Radial basis neural networks, Naïve Bayes, Decision trees J48, Support vector machines, and simple CART in recognizing illnesses to discover the best classifier [13]. Binary as well as multiclass datasets from UCI machine learning repository were utilized, namely Pima Indians Diabetes database, WBC, WDBC, and Breast tissue. The experiments were conducted in WEKA and the outcomes were compared and analyzed. It was established that SVM RBF Kernel outperforms others, when com-pared with the other algorithms in terms of accuracy, sensitivity, specificity and precision for both binary as well as multiclass datasets. Even though other classifiers performed well in classification applications, their behaviour differed for each dataset. In another study [14], ID3, Support Vector Machine (SVM), Naïve Bayes, C4.5, C5.0, K Nearest Neighbours (KNN) and CART were utilized. The various data mining algorithms were applied on Pima dataset, which consisted of 768 records of various patients, followed by their evaluation and comparison. Results demonstrated that the level of accuracy in SVM algorithm was about 81.77, and thus, it was more accurate. In addition, C5.0 decision tree, a standout amongst the most generally utilized methods was found to be more accurate when compared to ID3, C4.5 and CART models.

III. ARCHITECTURE

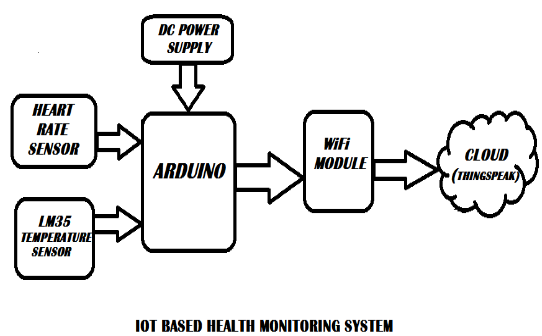

Fig 1. The Block Diagram of our system

Three-Layer Architecture:

A. Perception Layer

The perception layer was the physical layer in the system. This layer had sensors for sensing and gathering information about the environment, mainly a specific area of use. It sensed physical parameters or identified other smart objects in the environment. Parameters included in this layer were heart-beat or pulse, body temperature, etc.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Sensors like Temperature Sensor- LM35 Heartbeat. Sensor with Arduino was used in this layer to obtain the pulse of the patient.

B. Network Layer

The network layer was responsible for connecting to other smart things, network devices, and servers. Its features were also used for transmitting and processing sensor data.

We transferred the data obtained from the Arduino to the *ThingSpeak cloud* for storage and analysis.

- Network technology : wireless application.

C. Application Layer

The application layer was responsible for delivering application specific services to the user. It defined various applications in which the Internet of Things could be deployed, for example, smart homes, smart cities, and smart health, etc.

- Smart Health Monitoring.

IV. IMPLEMENTATION DETAILS AND HARDWARE INTERFACE

The implementation required Arduino Uno board, a pulse sensor, a temperature sensor (LM 35), Wi-Fi shield and a bread-board. The connections were made are as per the diagram below (Fig 2).

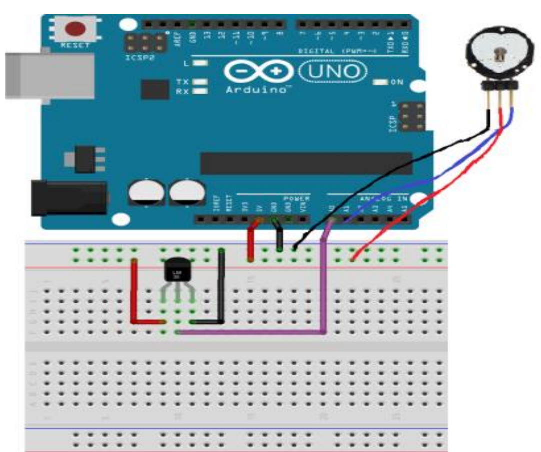

Fig 2. Diagram showing hardware interfacing between arduino and pulse sensor and temperature sensor

We used Arduino to help us interface the pulse sensor and The body temperature sensor to acquire data from patients and converted the analog inputs from the sensors into digital form. In order to have a more comprehensive detail, 3 sets of 300 readings (Dataset 1, Dataset 2, Dataset 3) from each patient were recorded over a period of two months. The main parameters considered were the heart rate and the body temperature of each of the individuals at various timings of the day.

A Wi-Fi shield was connected to the Arduino in order to Transfer data from Arduino to a cloud database. Here, we used *ThingSpeak* to store the data for further analysis after which we applied classification algorithms to all the three datasets.

We employed three classifiers namely -Naive Bayes, Random Forest and Support Vector Machine- for our model because of their high level of accuracy. The SVM model had the advantage that it was not influenced by outliers and was not sensitive to over-fitting. Naïve Bayes had proven to be efficient and like SVM, was not based on outliers. Since the dataset obtained was non-linear, these models worked well and followed a probabilistic approach. Lastly, Random Forest had been demonstrated to be powerful and accurate, providing good performance to various kinds of problems, especially on a non-linear problem set. We ran these classifiers on three different datasets obtained and recorded the results. The results were then plotted to analyze the variations in the applications of these machine learning models.

The size of the candidate training sets obtained are-

- Training set 1: 65
- Training set 2: 129
- Training set 3: 224

V. RESULTS AND DISCUSSIONS

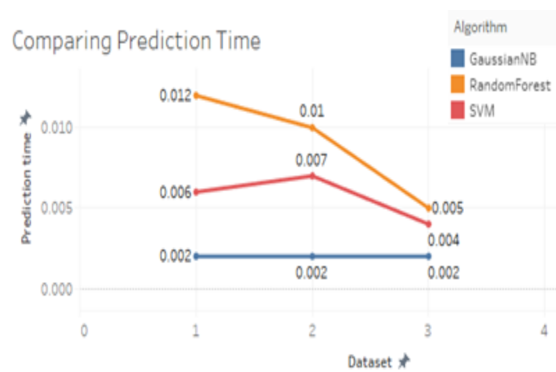

Fig 3. Comparison of prediction time of algorithms across datasets

Fig (3) shows that Naive Bayes is the fastest algorithm, followed by Random Forest and SVM.

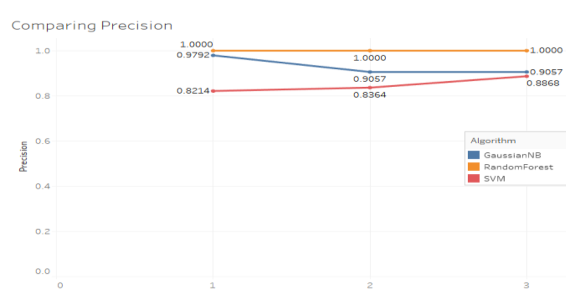

Fig4.Comparison of precision of algorithms across datasets

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Fig (4) shows that Random Forest has the best precision, followed by Naive Bayes and SVM. We can also notice that the precision of Naive Bayes decreases as dataset size increases while the precision of SVM increases with dataset size. It may so happen that on further increasing the dataset size, SVM may show better precision than Naive Bayes.

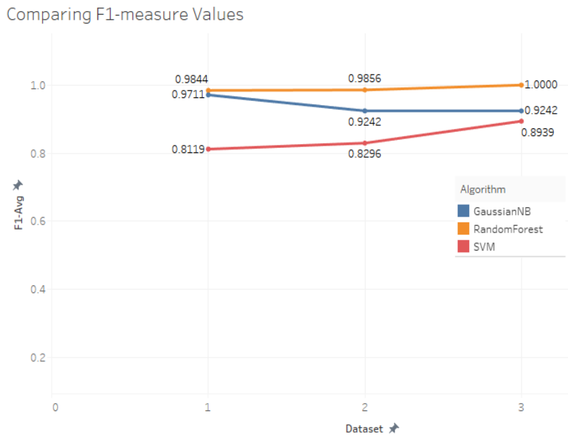

Fig 5. Comparison of F1-measures of different algorithms across datasets

F1-measure helped us identify the best classifier in terms of two parameters, namely, recall and precision and is defined as the harmonic mean of these factors. From Fig (5), the average F1 value of Random Forest is 0.9856, followed by Naive Bayes with value 0.9398 and then SVM with value 0.8451. Therefore, with respect to precision and recall, Random Forest is the best classifier.

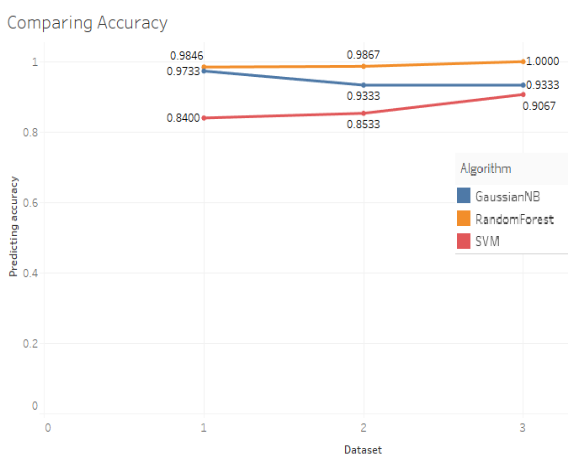

Fig 6. Comparison of accuracy of different algorithms across datasets

From Fig(6), we can see that Random Forest has the best accuracy followed by Naive Bayes and SVM.

VI. CONCLUSION

A novel method for health risk prediction based on heart rate and body temperature is outlined using comparison of three classifiers, namely, Random Forest, Naive Bayes and Support Vector Machine. Our experiment results showed that Naive Bayes was the fastest while SVM was the slowest classifier. With respect to precision and recall, Random Forest was the best classifier as it had the highest f1-score while SVM was the least preferable. The accuracy of Random Forest was highest while that of SVM was the lowest. Therefore, for health risk prediction, Random Forest was found to be the optimum choice.

REFERENCES

1. J. B. Rubin and W. B. Borden, "Coronary Heart Disease in Young Adults", Curr. Atheroscler Rep., vol 14, no. 2, (2012), pp. 140-149.
2. A. B. Donnersberger, "A Laboratory Textbook of Anatomy and Physiology", Jones and Bartlett Publishers, Illinois, (2009).
3. G E. Billman, "Heart Rate Variability-A Historical Perspective", Frontiers in Physiology., vol. 2, (2011), pp. 1-13.
4. E. Jovanov, A. Milenkovic, C. Otto and P. A. Groen, "A Wireless Body Area Network of Intelligent Motion Sensors for Computer Assisted Physical Rehabilitation", J. Neuroeng. Rehabil., vol. 2, (2005).
5. L. Eisman, "Biology and Human Progress", Englewood Cliffs, NJ: Prentice hall, (1972).
6. R. Trivedi, G. Mathur and A. Mathur, "A survey on Platinum Temperature Sensor", International Journal of Soft Computing and Engineering, vol. 1, (2011).
7. Rajvardhini Katake, Bhagyashree Kute, Sharmili Ranjane, Shubham C. Jaiswal, "Survey of Health Monitoring Management Using Internet of Things (IOT)", International Journal of Science and Research (IJSR), (2016).
8. Wagner, Matthias, *et al*., "Android-Based Body Area Network for the Evaluation of Medical Parameters", *Intelligent Solutions in Embedded Systems (WISES), 2012 Proceedings of the Tenth Workshop on* IEEE, (2012).
9. Hossein Fotouhi, Aida Causevic, Kristina Lundqvist and Mats Björkman, "Communication and Security in Health Monitoring Systems -- A Review", Computer Software and Applications Conference (COMPSAC), 2016 IEEE 40th Annual
10. Kirti Raghav, Jeet Paul and Raji Pandurangan, "Design and Development of e-Health Care Monitoring System", International Journal of Research (IJR), vol 2, no. 4, (2015), pp. 977-981.
11. Mohammed, Junaid, *et al*., "Internet of Things: Remote Patient Monitoring Using Web Services and Cloud Computing", 2014 IEEE International Conference on Internet of Things (iThings), and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom), (2014).
12. Kaiser, M.S., Paul, M., Rahman, M.M., Reza, S.M. and Sarkar, S., "Low Cost and Portable Patient Monitoring System for e-Health Services in Bangladesh", *International Conference on Computer Communication and Informatics (ICCCI)*, 1-4. (2016).
13. S. Aruna, S.P. Rajagopalan and L.V. Nandakishore, "An Empirical Comparison of Supervised Learning Algorithms in Disease Detection", International Journal of Information Technology Convergence and Services (IJITCS) Vol.1, No.4, (2011).
14. Farahmandian, Masoum, Yaghoub Lotfi, and Isa Maleki, "Data Mining Algorithms Application in Diabetes Disease Diagnosis: A Case Study", *Technical Report*. MAGNT Research Report, (2015).

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication