Common Disease Proneness Prediction Using IOT and Artificial Intelligence

1. Karan Ahir, computer dept, DY. Patil School of Engineering Academy,Ambi
2. Deepali Jadhav, computer dept, DY. Patil School of Engineering Academy,Ambi
3. Devendra Naik, computer dept, DY. Patil School of Engineering Academy,Ambi

Abstract

Monitoring of various health parameters everyday is not viable as it requires the investment of a lot of time and effort. Traveling all the way to the clinic for a doctor's appointment daily is quite expensive and time-consuming. As the Vital signs and other parameters can be monitored by various gadgets, such as wearable sensors and fitness trackers as well as smartphones nowadays. Most of their gadgets are small and portable in nature as they are supposed to be worn or carried without any inconvenience, which in turn reduces the battery capacity as well as the memory of the device to fit in such a form factor, which limits the scope of Healthcare predictions. Therefore, to ameliorate this effect, the researchers have implemented a technique that utilizes the IoT platform to outsource the recorded data on to the server. The server employs K Nearest Neighbour and Artificial Neural Network that is capable of accurately predicting the Health conditions of a person with the help of parameters, such as Pulse Rate, Blood sugar level, body temperature, etc.

Keywords: *IOT, ANN, K-Means,Rule based classification.*

I. INTRODUCTION

Healthcare is one of the most important aspects of a satisfactory and fulfilling life. Healthcare has facilitated a lot of aspects of human life that were almost non-existent when it wasn't as pronounced as it is today. Back in the Middle Ages, there was rampant death of infants and their mothers during childbirth, the life expectancy was at an all-time low. As we advanced through the generations, our understanding of Healthcare also grew with the average life expectancy and child mortality.

This is because humans have been actively researching into the field of medical science to ascertain cures of various ailments affecting the human body as well as finding cost-effective alternatives to expensive medications that can be beneficial and would allow for a ubiquitous presence. Healthcare revolves around very basic paradigms, such as the physical and psychological wellbeing of the people. This is the motto of the healthcare institutions that provide healthcare for the patients.

Internet of Things

Due to technological advancements, there has been exponential growth in the healthcare sector. As technology advances, it is supposed to provide a greater degree of comfort and convenience to its users, which is one of the widely publicized benefits of the Internet of Things Technology. Internet of Things is one of the most important aspects of technology that can readily be employed in the Healthcare sector and will be capable of reaping huge benefits from this platform.

The Internet of Things platform is designed to be able to connect almost all the electronic and electrical appliances and components in a house or an office etc to the internet. This is a very innovative concept that can allow for seamless communication and open up doors to various techniques that would benefit the users of this technology. As most of the devices in the house will be connected with each other and the internet, it would be highly beneficial to the user as it can allow for greater control while increasing the convenience.

Internet of Things has been widely used in the field of Medical science as a way to achieve almost ubiquitous coverage for the purpose of almost constant monitoring of the patient. this is highly useful for early discharge of patients and also to monitor the health of a critically ill patient or an elderly patient from the comfortable confines of their house. This is very useful to allow for the convenience of the patients as well as the doctors who have very limited free time from their work schedule.

Artificial Neural Network is one of the most commonly used techniques of Machine Learning. Due to the fact that Artificial Neural Networks are quite easy to deploy and manage. Artificial Neural Network derivesits name from the Neural Networks that exist in human brains as they are modeled after the inner workings of the human brain. Both, the human brain as well as the Artificial Neural Network has their basic unit as the neuron.

There are billions of neurons that make up the human brain. These Neurons are utilized in the Artificial Neural Network in the same way they work in the brain. This helps in creating a network that is capable of behaving and reacting in a similar way a human would react. This is highly useful in various applications such as, Signal processing, Telecommunication, telecom, medical fields, etc. the Artificial Neural Networks are capable of performing tasks that are difficult for a machine to perform but are easier for a human to execute the said task efficiently.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

II. LITERATURE SURVEY

K.Navinprovides the solution for the disease outbreaks for the public health epidemic surveillance system. There is a need to improve modern technology adaptions as there is growth in population and the various infectious disease are increasing it is a challenge to the public health sector.[1] The very first testing ofsmart epidemic surveillance was done by using the cloud push and the real-time database and the notification was fused with mobile as it is used as the backend services. The system provides fast and ordered data communication

Arthi C. introduces the problems of the teachers undergo such as the illness and health hazards. The teaching profession is assumed as the easiest and the safest professions with less physical strain. There isa very small number of systems which predict profession related illness they only focus mainly on field related issues.[2] The proposed system focuses on the disease such as voice disorders, chronic laryngitis and respiratory illness like asthma that affects because of their profession.This technique could help teachers in earlier detection of disease.

Y.Jung states that there is a high demand for physical and mental illness for improving the quality of life.[3] In this paper monitoring of the disease done by using themultimodal biosensors the sensors such as EEG (electroencephalography), ECG (electrocardiography), respiration rate, SpO2 and skin temperature. From this sensor, we get the measurement that is used to generate the bio-emotional index. To analysis, this Expectation Maximization (EM) is used for health status prediction

S.Kesavan[4] describes the gateway as a solution that connects the devices or sensor to the IT network and cloud services because of these there is communication between the world and devices and network there are also other wireless technology such as ZigBee, Bluetooth, Z-wave used in low rate data communication and the Wi-Fi is used in high rate data communication. By the occurrence of two or more things coming together of digital technology thesmart devices, the dream for monitoring and tracking data in real time become true. Use of sensor has made smart devices, monitoring the health become easy.

N.Alshurafa elaborates remote health monitoring becoming the most widely hospital to monitor and communicate with the patients in the given time and by reducing hospital cost and the quality of treatment Wanda-CVD is a smartphone depends on a RHM system designed to assist participants in decreasing identified cardiovascular disease (CVD) risk factors through wireless coaching using feedback and prompts as social support.[5]This Wanda-CVD was developed by the Women's Heart Health Study (WHHS).

Y.Jung explains the mobile application becomes of the smart device which provides health services anywhere anytime due to the growth of the mobile wireless services. [6] To provide the health service there is need of the middleware which provides the inner information to the outside world these sensors are EEG (electroencephalography), ECG (electrocardiography) and EMG (Electrocardiogram) waveforms from wearable ECG devices comes under the coverage of a wireless sensor network (WSN).

R.Kukafkadevelops Therapeutic Lifestyle Change Decision Aid (TLC DA) system to support the behavior change to work when there are multiple unhealthy behaviors are present. The information is gathered from the system to generate the modified messages to consumers coerce them into following certain healthy lifestyles. [7]The main aim of the paper is to study and find primary determinants of health behavior choices made by patients by using TLC DA system. The parameters include smoking status, smoking cessation success estimate, self-efficacy, body mass index and diet status.Prediction of smoking cessation choice was the most accurate (73%) and weight management choice (67%).

A.Yassinedescribes that there is vast growth in the population due to people relocating from village to city. Health care is one of the important factors to provide people in the city. [8] There is vast investment by the client from around the world and from the government. There are millions of home equipped with smart devices and the sensor so they analyses and Indexes data

to support the smart city services.The results of identifying human activity patterns from appliance usage are presented in detail in this paper along with the accuracy of short and long term predictions.

Z.YANstates mobile smart terminal has become one of the crucial and important parts of the people lives. There has been a great development in the field ofelectronic technology and in biological technology which makes possible to get the physical information detection by using the healthy smart terminal.[9]This terminal is very strong in the computational processing ability, cloud data storage,communication ability,and powerful data mining ability. The core part of the project is based on user motion information, emotion, pressure, and other personal information, disease information, weather, climate, pollution, environment, influential event information.The results will provide a reference for disease prediction and early warning analysis.

D.Kellydevelops a sensor based methodology to measure health status.[10] There are different techniques used to generate the human behaviors that can be done from the smartphonesaccelerometer and gyroscope sensors. The measurements and the output from these devices are given to system the output from the sensor evaluated and used to predict health status. Using the proposed, method the health status can be measured using the modest and already available hardware.

T.Brisimi explains the modern living in the big cities and the side effects on health in recent years there is vast growth inchronic diseases.[11] Two leading chronic diseases are heart disease and diabetes. Thus, living in the city is impacting health in a different ways such as stress associated with fast-paced urban life, a sedentary lifestyle due to work conditions and lack of time, air pollution, and a disproportionate number of people living in poverty, urban populations face an increased risk for the development of chronic health conditions. They have to developtheoretical out-of-sample guarantees for the latter method.

A.Najar introduces one of the major problems in Indonesia that is Dengue Hemorrhagic Fever (DHF). Due to increasing population weather changes, epidemic factors there is an increase in the growth of dengue fever patient. To avoid this outbreak it is necessary to apply the extreme learning machine (ELM) method to predict the risk of the outbreak based on weather conditions.[12] They have used binary sigmoid activation functionand bipolar sigmoid with number hidden from 5-200 nodes. In the future, this model can be compared with SVM for prediction

K.Dolui [13] explores novel services that are used to Real-Time Health Advice and Action (ReTiHA) by monitoring the patient parameter and medical advice. The main aim of the paper is to provide health services in the remote area where the health system is unavailable. They have used wearable sensor and the smart devices to collect the data from the patient end and i=o the basis of that monitoring of the patient is done. With the advent of the Internet of Things (IoT), many newcomponents can be added to the current system.

III PROPOSED METHODOLOGY

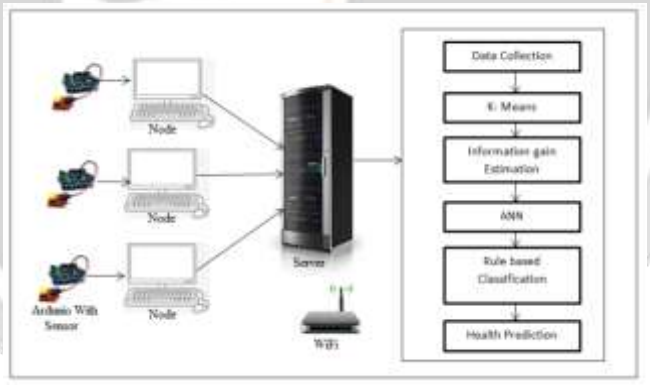

Figure 1: Overview of the Proposed model

The above Figure 1 depicts the flow of the presented Healthcare system that employs the Internet of Things. The various methodologies deployed in this technique have been elaborated below.

Step 1: Deployment and Data Generation – A distributed System has been utilized for the deployment of the healthcare system, which includes a panel for the registration of the patient who is slated to use the Healthcare prediction service. The patient has to provide all the basic details such as weight, height, diabetic, smoker, etc. After providing the details the patient is asked to wear the pulse rate monitor to start the service.

The blood sugar levels, boy temperature, blood pressure, etc are calculated based on the equations provided by the famous website for medical services, www.webmd.com. The data generated from these random integer protocols is then transferred to the prediction server, which has been set up in a LAN connection with the help of Wifi.

Step 2: Data Pre-processing – As soon as the prediction server receives the patient data from the patient panel in the distributed system, the received data is then analyzed for that specific patient by pulling their data from the database. If the data received is not sufficient for the analysis, the data is then stored to collect subsequently more data.

If the data is sufficient, then the extracted data is pre-processed by subjecting it to collecting only the changing data in the stipulated time interval. The data that is selected for pre-processing includes Sleep hours, Blood pressure, Blood sugar, body temperature, and Pulse rate.

Step 3: K means Clustering – The Euclidean distance is calculated for the pre-processed data for each of the rows in accordance with the other rows which is then referred to as the R_D. The pre-processed data is then clustered according to this Euclidean distance. The complete Euclidean Distance (E_D) of all the rows is evaluated from the average of all the rows combined.

The generated E_D is then utilized as a centroid with the boundary. The randomly picked centroids correspond to the values of K and the R_D which is the row distance. The boundaries are again decided with the help of the centroids until the desired size of the cluster is reached. The Equation for the Euclidean Distance is as follows.

$$\boldsymbol{D = \sqrt{(x1 - x2)^2 + (y1 - y2)^2}} \text{ ---------- (1)}$$

Where,
D- Euclidean distance
x1,x2,y1, and y2 are the attribute values

Step 3: Information Gain Estimation – The information gain theory helps identify the important clusters from all the generated ones in the previous step. The serious health conditions that have the potential to be highly dangerous are recognized from the pre-processed data. The Shannon and Gain theory are then applied to the rows that exhibit 75% threshold of the serious health conditions in equation 2.

$$\text{IG} = -\frac{P}{T}\log\frac{P}{T} - \frac{N}{T}\log\frac{N}{T} \text{ \_\_\_\_\_(2)}$$

Where

P= Frequency of the Ill health parameter count
T= Cluster Elements Size.
N= T-P
IG = Information Gain of the cluster

Step 3: ANN – The mean and the standard deviation of the R_D are calculated and then information gain theory is used to select clusters to form the neurons on the Artificial Neural Network. The protocols governing the neurons are the standard deviation and mean that has been calculated. The workings of the Artificial Neural Network in this context is depicted in the Algorithm below.

Algorithm 1: Neuron formation

// Input : Selected Clusters C_L
// Output : Neurons N_U
Function :neuronFormation($\mathbf{C_L}$)
Step 0: Start
Step 1: N_U =Ø
Step 2: ***for***i=0 **to** size of C_L
Step 3: $S_G = C_{Li}$
Step 4: mean and Standard _deviation for S_G as (μ,α)
Step 5: ***for*** j=0 **to** size of S_G
Step 6: ROW=S_{Gj}
Step 7: R_D = ROW(ROW_{SIZE} -1)
Step 8: **If**(RD> = (μ- α) AND $\mathbf{R_D}$< =(μ+ α)
Step 9:$N_U=N_U+ROW$
Step 10: ***ELSE***

Step 11: ***If*** $R_D > \mu + \alpha$, **THEN**
Step 12: $N_U = N_U + ROW$
Step 13: ***ELSE***
Step 14: ***If*** $R_D < \mu - \alpha$, ***THEN***
Step 15: $N_U = N_U + ROW$
Step 16: **End *for***
Step 17: **End *for***
Step 18: return N_U
Step 19: Stop

Step 4: Rule-based Classification – The Neurons generated in the penultimate step have a tendency to evaluate the prediction of diseases based on the variation of the parameters that govern the formation of the neuron. These neurons are further utilized to evaluate the rest of the patient data that is being extracted.

When all of the data is analyzed by the Artificial neural network, it is then segregated binarily based on the factors, such as Alcohol consumption, Blood Sugar levels, Smoking Habits, body temperature, and pulse rate.

Therefore, the output of this binarization is again checked and counted for the evaluation of the various ailments the person might be suffering or prone to such as Diabetes, Anxiety, Insomnia, Agoraphobia, Fever, etc. if there are no correlations between the patient data and the diseases, the patient is considered fit and healthy and is intimated of the same via an Email.

IV RESULT AND DISCUSSIONS

The Disease prediction approach has been successfully implemented in a Local Area Network with the help of a D-Link router with dual antennas which was used for the wireless communication purposes. The system for the proposed methodology was deployed on a machine with an average configuration of a core i5 Central Processing unit coupled with 6GBs of RAM. Arduino UNO microcontroller has been utilized for the patient panel for the purpose of recording the pulse rate of the patients through the connected sensor. NetBeans IDE was used for the Java programming implementation and the database services were effectively handled by MySQL.

Root mean square Error (RMSE) is used to measure the error rate of the prediciton system. Here in this experiment a error rate is measured in between the actual Diseases and predicted disease using RMSE.And this RMSE can be depicted using the equation given below.

$$\mathrm{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

Where
$\sum$ - Summation
$(Zfi - Z_{oi})^2$ - Differences Squared for the Smma
N - Number of samples or Trails

No of Trails Conducted	Actual Diseases	Predicted Diseases	MSE
5	4	4	0
10	5	5	0
15	8	6	4
20	11	8	9
25	13	9	16

Table 1: Mean Square Eror Reading

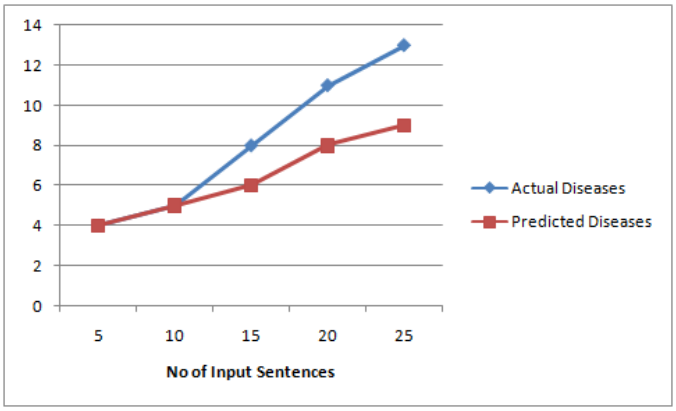

Figure 2: Mean Square Error (MSE) for the Actual and Predicted Disease.

The above experiment yields average MSE of 5.8 and RMSE of 2.4. This indicates the error rate of the proposed model is too low in prediction of the common diseases. So this is a better sign of any prediction system and it also shows the effectiveness of the proposed algorithms.

V CONCLUSION AND FUTURES SCOPE

The presented approach for utilizing Machine Learning and IoT for the purpose of predicting ill health conditions in an individual is successfully implemented with the help of the random attribute generation. The real-time data consisting of the patient's pulse rate has been utilized effectively in accordance with extensive medical protocols to provide valuable insight. Machine Learning algorithms such as K-means Clustering and Artificial Neural network with the addition of Rule-based classification has performed significantly better than the traditional technique which is evident from the precision and recall values calculated.

For the purpose of future research, EEG and ECG signals can be added as parameters with the help of IoT sensors in real time to achieve even greater precision.

REFERENCES

[1] Mr. K.Navin, Dr.M.B. Mukesh Krishnan, Mrs. S.Lavanya,Dr.A.Shanthini,"A Mobile Health Based Smart Hybrid Epidemic Surveillance System To Support Epidemic Control Programme In Public Health Informatics "International Journal of Computer Applications, Vol 61, No. 21, pp. 45-50, 2017.

[2] Arthi C. Priya R., Ruchi Rautela,"Analysis and Prediction of health issues for teaching profession using Semantic Techniques "2013 International Conference on Informatics, Electronics & Vision (ICIEV), IEEE Xplore., 2017.

[3] Yuchae Jung,Yong Ik Yoon,"Wellness Contents Recommendation Based on Human Emotional and Health Status Using EM "978-1-4799-8993-5/15/$31.00 , IEEE,2015.

[4] Selvaraj Kesavan,Ganesh Kumar Kalambettu," IOT enabled comprehensive, plug and play gateway framework for smart health"2018 Second International Conference on Advances in Electronics, Computer, and Communications ,CAECC-2018

[5]Nabil Alshurafa, Costas Sideri," Remote Health Monitoring Outcome Success Prediction using Baseline and First Month Intervention Data"IEEE Journal of Biomedical and Health Informatics. ,2016.

[6]Yuchae Jung,Yongik Yoon,"Prediction Model for Mental and Physical Health Condition using Risk Ratio EM "978-1-4799-8342-1/15/$31.00 ©2015 IEEE 439 ICOIN ,2015.

[7]Rita Kukafka,Cheol Jeong, Joseph Finkelstein," Introducing Decision Support for Smart Mobile Health Behavior Change Applications"978-1-4799-7303-3/15,IEE ,2015.

[8]Abdulsalam Yassine, Shailendra Singh, Atif Alamri,"Abdulsalam Yassine, Shailendra Singh, Atif Alamri"in 2015 International Joint Conference on Neural Networks (IJCNN), pp. 1–7, July 2015.

[9]Zheng-hangYAN,Yang LIU,"The Multidirectional Information Fusion Mobile Health Management Technology"978-1-4799-0766-3/13, IEEE,2018.

[10]Daniel Kelley, Kevin Curran, Brian Caulfield."Automatic Prediction of Health Status using Smartphone Derived BehaviourProfiles "IEEE Journal Of Biomedical And Health Informatics..2016

[11]Theodora S. Brisimi, Tingting Xu, Taiyao Wang, Wuyang Dai, William G. Adams,"Predicting Chronic Disease Hospitalizations from Electronic Health Records: An Interpretable Classification Approach"0018-9219 © IEEE.,2018

[12]Abdul MahatirNajar,Mohammad Isa Irwan,DiekyAdzkiya,"Extreme Learning Machine Method for Dengue Hemorrhagic Fever Outbreak Risk Level Prediction "International Conference On Smart Computing And Electronic Enterprise (ICSCEE2018) ,IEEE,2018.

[13]KoustabhDoluiSrijani Mukherjee, Soumya Kanti Datta, VenkiteswaranRajamani,"ReTiHA: Real Time Health Advice and Action using Smart Devices " International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICC) , 2014

*****