International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-10, Oct 2018 E-ISSN: 2347-2693

An Intelligent Automatic Multi-Disease Prediction Technique using Data Mining Algorithms and Big Data

[1*]S Manimekalai, [2]R Suguna, [3]S Arulselvarani

[1,2]Dept. of Computer Science, Theivanai Ammal College for Women (Autonomous),Villupuram.
[3]Dept. of Computer Science, STET Women's College, Mannargudi.

[*]*Corresponding Author: mamekaroshni@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 15/Oct/2018, Published: 31/Oct/2018

Abstract - Big data and its techniques not only help the biomedical and healthcare sectors to forecast the disease prediction but also the patients. It is difficult to meet the doctor at all the times in hospital for minor symptoms. Big data gives necessary information about the diseases based on the symptoms of the patient. Nowadays people wants to know more about their health, diseases and the related treatments for their betterment. However existing health care system gives structured input which lacks in reliable and accurate prediction. Here, Automatic Multi-Disease Prediction (AMDP) technique is proposed which identifies the most accurate disease based on patient's input which benefits in early detection. Electronic Health Record (EHR) maintains and updates patient health records which facilitate an improved prediction model. Big data uses both structured and unstructured inputs which result in instant guidance to their health issues. The system takes input from the users which checks for various diseases associated with the symptoms based upon analyzing a variety of datasets. If the system is not able to provide suitable results, it intimate the users to go for Clinical Lab Test (CLT) such as blood test, x-ray, and scan so on where the uploaded images are sent for the effective deep learning prediction. The different parameters included in effective automatic multi disease prediction include preprocessing, clustering and predictive analysis. The main objective of the proposed system is to identify the diseases based on the symptoms and give proper guidance for the patients to take treatment quickly without making any further delay in a convenient and efficient manner.

Keywords - Big Data, AMDP, EHR, Deep Learning Algorithm, CLT.

I. INTRODUCTION

Due to the contribution of internet technology makes our life more convenient and provides a variety of information from which one can browse their desired needs. Big data has shown progressive hike transformation in medical and health care sectors [11]. It provides analytical methods for collecting valuable information, maintains and updates individual health profile and various prediction models. Through advanced technologies towards which make use of all clinical data and achieves better disease prediction in order to treat various disease and provides clinical decision support[19].

Revolution of internet plays a major role in healthcare information systems containing huge number of medical records are ideal targets for data mining [18]. But it can't possibly happen that everything cannot be at their fingertips. The main challenge is to provide efficient and wealthy information relevant to the search. The detailed research and statistical analysis is more than what a human mind can think. It is the reason everyone is moving towards predictive analytics [13].

Predicting disease is essential for any medical organizations in order to make the best possible medical care decisions result. Ithunts through massive number ofinformation, and analyze outcomes for users [17]. It also looks data from past results as well as latest advancements. Incorrect decisions may cause delays in medical treatment or even loss of life. The main objective is to predict the diseases from the given symptoms create and monitors a health profile of every individual patient [7].

With the proliferation of medical data, collecting electronic health records (EHR) contains patient's information; test results and patient history. In order to predict disease sever factors has been consider such as body mass index, cholesterol level, blood sugar, blood pressure and so on. In contrast, our proposed predict an accurate disease based on other diseases that a patient already has. One of the benefits of these approaches is that it can be applied to predict anydisease rather than a specific disease [2].

Existing system uses only structured data so it's difficult to predict the accurate disease prediction. Many of the existing system focus only on particular disease. In order to overcome this issue, the proposed system propounds an

accurate automatic multi-disease prediction in health care using big data [17]. It creates an automatic system to diagnose the disease based on the user symptoms and their previous health history. Furthermore, compares with clinical laboratory test report due to higher complexity. Lastly decision tree provides an accurate result. Through this patient can easily diagnose their disease by just entering their issues or symptoms and provides result about what he or she suffering from[14].

Effective analysis of an accurate automatic multi-disease prediction techniques include like preprocessing, clustering, frequent pattern growth algorithm, pattern learning algorithm [15], Naive bayes classifier, Profile based analysis, electronic health record, clinical laboratory test and decision tree. The overall aim of this proposed system is to make the suitable treatment quickly and easily [18].

The paper is organized as follows: Chapter II discussed about the review of the literature about this field. Chapter III gives the proposed methodology. Results and its discussions are presented in Chapter IV. Conclusions are given in Chapter V.

II. LITERATURE REVIEW

Min Chen, et.al [1] proposes machine learning algorithms for effective prediction of chronicdisease outbreak in disease-frequent communities. Experimentthe modified prediction models over real-life hospital data collected from central China in 2013-2015. To overcome the difficultyof incomplete datauses a latent factor model to reconstructthe missing data on a regional chronic diseaseof cerebral infarction. Moreover, different regions exhibit uniquecharacteristics of certain regional diseases, which may weakenthe prediction of disease outbreaks.Proposes a new convolution neuralnetwork based multimodal disease risk prediction (CNN-MDRP) algorithm using structured and unstructured data from hospital. Here focusedon both types of structured and unstructured data in the area of medical big data analytics.Compared to several typical prediction algorithms, the predictionaccuracy of our proposed algorithm reaches 94.8% with aconvergence speed which is faster than that of the CNN-basedunimodal disease risk prediction (CNN-UDRP) algorithm.

Feixiang Huang, et.al [2] applies the data mining process to predicthypertension from patient medical records with eight otherdiseases. A sample with the size of 9862 cases has been studied.The sample was extracted from a real world Health care Information System database containing 309383 medical records. Here, observed that the distribution of patient diseases in themedical database is imbalanced. Under-sampling technique has been applied to generate training data sets, and data mining Weka tool has been used to generate the Naive Bayesian and J-48 classifiers. In addition, an ensemble of J-48 classifiers was created trying to improve the prediction performance, and roughset tools were used to reduce the ensemble based on the idea ofsecond-order approximation. Experimental results showed a littleimprovement of the ensemble approach over pure Naïve Bayesian and J-48 in accuracy, sensitivity, and F-measure.

Sujatha R, et.al [3] enhances mining in health sector plays an important role for uncovering new trends which in turn helpful for all the parties associated with this field. It is a new powerful technology which is of high interest in computer world. It is a sub field of computer science that uses already existing data in different databases to transform it into new researches and results. It makes use of Artificial Intelligence, Machine Learning and database management to extract new patterns from large data sets and the knowledge associated with these patterns. The actual task is to extract data by automatic or semi-automatic means. The different parameters included in data mining include clustering, forecasting, path analysis and predictive analysis.

Pinky Saikia Dutta, et.al [4] describes about data mining in disease prediction is a sub field of computer science that uses already existing data in different databases to transform it into new researches and results. In this data-rich world, people are running out of information. This can be a matter of risk for the person who needs immediate remedies regarding their poor health. To unfold this hurdle, the concept of data mining is the best suited. Here, the traditional approaches have been replaced by smart technologies. The main purpose of data mining application in healthcare system is to develop an automated tool for identifying and disseminating relevant healthcare information. In this system, we have presented a web based application for Predicting diseases based on user input symptoms. It predicts probable diseases by mining data sets and provides remedial solutions for effective treatment.

Ravi Aavula, et.al [5] suggest e-health applications because consulting a doctor is a quite obvious thing in our day to day life, but the availability of the doctor during the time of our requirement is unpredictable. In order to overcome this issue a proposal of android application is made, this smart health application enables users to get instant report on their health issues through an intelligent health care application online. This e-health application enables user to express their symptoms and issues. It then processes user's issues and symptoms to check for various health issues that could be associated with the symptoms given by the user. If the application is unable to provide a particular solution then it urges the user to under-go tests like blood test, CITI scan accordingly.

Priyanka Vijay Pawar, et.al [6] describes about the importance of smart technologies over traditional approaches has been integrated in every discipline of science. Here, propounds a methodology to predict diseases based on user input symptoms. They have built a prototype to demonstrate the efficiency of these methods which will inform users about the disease they are suffering from. It predicts probable diseases by mining data sets and provides suggested doctors and remedial solutions. It will also guide the users by giving tips to live a healthy life, some diet tips and also usefulness of plants and food items. The authors identified probability of diseases using apriori technique.

Krishnaiah V, et.al [7] illustrates an overview of the applications of data mining techniques, medical, research, and educational aspects of clinical predictions. In medical and health care areas, due to regulations and due to the availability of computers, a large amount of data is becoming available. On the one hand, practitioners are expected to use all this data in their work but, at the same time, such a large amount of data cannot be processed by humans in a short time to make diagnosis, prognosis and treatment schedules. A major objective of this paper is to evaluate data mining techniques in clinical and health care applications to develop accurate decisions. The paper also provides a detailed discussion of medical data mining techniques can improve various aspects of Clinical Predictions.

Rawan Ali, et.al [8] creates an online consultation project for health prediction system. It allows users to get instant guidance on their health issues through an intelligent health care system. Through onlinecontains data of various symptoms and the disease/illness associated with those symptoms. It also has an option for users of sharing their symptoms and issues. Data mining technique is used for prediction for patient disease and patient can take the medicine directly from the pharmacy if doctor can't make prediction.With the use of big data in biomedical and healthcare communities, accurate analysis of medical data benefits early disease detection, patient care and community services. However, the analysis accuracy is reduced when the quality of medical data is incomplete. Moreover, different regions exhibit unique characteristics of certain regional diseases, which may weaken the prediction of disease outbreaks. To overcome the difficulty of incomplete data, uses a latent factor model to reconstruct the missing data. A convolutional neural network based multimodal disease risk prediction (CNN-MDRP) algorithm using structured and unstructured data from hospital. To the best of our knowledge, none of the existing work focused on both data types in the area of medical big data analytics. Compared to several typical prediction algorithms, the prediction accuracy of our proposed algorithm reaches 94.8% with a convergence speed which is faster than that of the CNN-based unimodal disease risk prediction (CNN-UDRP) algorithm.

Md. Tahmid Rahman Laskar, et.al [9] propose ADPS due to rapid proliferation of internet technology and handheld devices has opened up new avenues for online healthcare system. There are instances where online medical help or healthcare advice is easier or faster to grasp than real world help. People often feel reluctant to go to hospital or physician on minor symptoms. However, in many cases, these minor symptoms may trigger major health hazards. As online health advice is easily reachable, it can be a great head start forusers. Moreover, existing online health care systems suffer from lack of reliability and accuracy. Herein, we propose an automated disease prediction system (ADPS) that relies on guided user input. The system takes input from the user and provides a list of probable diseases. The accuracy of ADPS has been evaluated extensively. It ensured an average of 14.35% higher accuracy in comparison with the existing solution.

III. PROPOSED METHODOLOGY

Health is a most important factor for everyone. But unfortunately it has been neglected today for many reasons. Absence of doctor due to some reason during emergency situation may result in loss of life. Not only that sometimes patient often feel hesitant to go to hospital for minor symptoms. These may prompt into major illness.

Big data mining enables distributed storage and processing environment is essential to store and process the healthcare data, which can be accessed at anytime. With the proliferation of technology in health care becomes easier to diagnosis any disease. It is applied in healthcare to identify the clusters of patients, diseases and future predictions with the help of various machine learning tools.

So, this paper proposed an automatic system that allows users to get instant guidance regarding their health issues. In order to diagnose any disease doctor initially analyses the symptoms of the patient after then final result is predicted. Similarly, machine diagnosis the diseases based on the symptoms just like the doctor does.

The system is fetched with various symptoms and their disease related with it. The user share their symptoms and issues. The health data are attributed as big data, which is defined by 3Vs in terms of Volume, Velocity and Variety. The collected patient data are of peta or zeta bytes, which describe the volume. The velocity is expressed in terms of data arrival rate from the patients. Variety explains the diversified data sets with respect to the structured, semi-structured and unstructured data sets such as scanned images, clinical reports, EHRs, and clinical lab test.

MapReduce framework collects data which are transformed into valuable insights.

By using intelligent data mining techniques evaluates the various disease related with it and predict most accurate disease. The system should respond like a doctor. The proposed system not aims to replace the doctor instead it helps to predict the disease and get proper medical assistance from the specialist. Along with the symptoms given by the patients the proposed system also using patient's history to make our results better. The proposed system uses Apriori-Frequent Pattern (A-FP) data mining algorithm to make these predictions almost perfect.

The A-FP algorithm is to find diseases based on the input symptoms through medical data mining. This algorithm is used to find medical datasets from which association rules can be generated. The aim is to find appropriate and frequent diseases from the medical dataset. As per Apriori a subject of a frequent itemset should also be a frequent itemset.

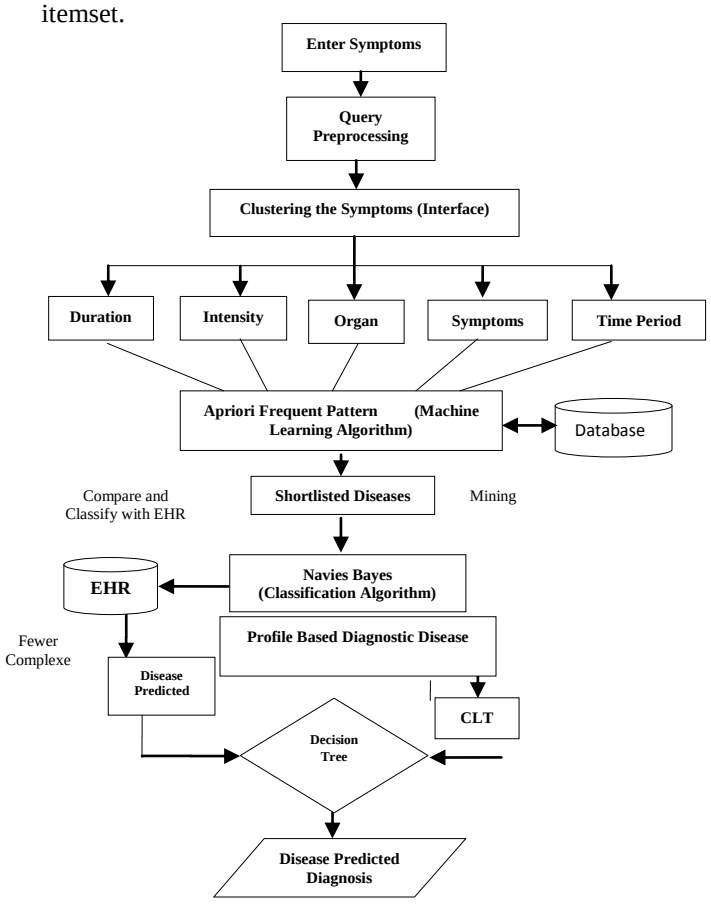

Figure 1

The shortlisted disease has been get from apriori frequent pattern algorithm. After this process, the patient data are combined with Electronic Health Record (EHR) to provide reports for better suggestions and decisions using naives' bayes algorithm. EHR includes doctor prescription, clinical reports and any previous clinical history of the patients. Even, the existing models cannot support both analysis and processing for the large volume of multi-structured healthcare data.

Naives Bayes aim to improve disease treatment and also its diagnosis in early stages for a faster and better treatment. Therefore, any NBNs attempt to make a faster and more accurate disease prediction and help the physicians for making a reliable decision. Let D be a training set of database. Each database is represented by an n-dimensional attribute vector. y includes 'n' independent attributes (y_1, y_2, …, y_n). Suppose there are 'm' classes such as c_1, c_2, …, c_m, then classification is to derive the maximal $p(c_i|y)$. This can be derived from Bayes' theorem as equation (1),

$$p(c_i\ |y) = p\ (y\ |\ c_i).\ p\ (\ c_i) / p(y) \qquad (1)$$

p(y) needs to be maximized because it has equal value for all classes as equation (2)

$$p(c_i\ |\ y) = p(y|c_i).p(c_i) \qquad (2)$$

An easy assumption in NBN is that the attributes are conditionally independent. So, the class assignments of this test samples are according to equations (3) and (4);

$$p(y|c_i) = \sum_{k-1}^{n} p(y|c) \qquad (3)$$

$$arg\ max\ c_i\ \{p(y|c_i).p(c_i)\} \qquad (4)$$

Bayesian logic can show the result of a patient's test, to predict or determine the chance of finding a disease. Bayesian theory implies that a rule for inferring or updating the amount of 'belief' in the light of new information. Bayesian networks can be seen as a substitute for logistic regression models that can formulate dependency or independency of variables.

Profile based diagnostic disease approach in medical field suggests based on their profile and prescripts according to user is an issue. Hence, necessary parameters such as, age, body mass index, so on and considering current body condition of the patient and other necessary details will improve in taking decision in optimal way. It leads to increases the quality of service provided to the patients and decreases the medical costs.

In case of high complex disease identified, the system is not able to provide accurate results, it urges users to go for Clinical Laboratory Test (CLT) or whichever report it feels user's symptoms are associated with. Most of the existing system not focused on unstructured input. By using convolutional neural network, both structured and unstructured data can be evaluated such as blood test, x-ray or any scan.

Finally, by using decision tree provides a most accurate result for the respective user/patient. Decision tree learning

uses a predictive model which maps observes about an item to concludes about the item's target value. It is one of the predictive modelling approaches used in statistics, data mining and machine learning. In these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. The proposed system predicts the accurate disease based on the symptoms and entered by the user. By considering user's previous medical history and profile based analysis to predict the most possible disease.

IV. RESULT AND DISCUSSION

The result of experimental analysis in identifying important patterns for predicting various diseases are presented. The systemis preprocessed effectively by removing related records and given that missing values. The well-mannered OWL dataset, resulting from preprocessing, A-FP mining, NB is then composed by decision tree. The performance analysis was evaluated by using performance indices such as sensitivity, specificity, PTR, NTR and accuracy for 3 patients. Some of the main formulations are as follows:

Accuracy: The accuracy represents the total accuracy rate of classifying each subject into the correct group. This index not only represents the probability of accurately classifying the subject as healthy or not, but also correctly classifying each patient into the correct disease group

Accuracy = (TP+TN)/ (TP+FP+TN+FN) * 100 %

Where TP is True Positive

TN is True Negative

FP is False Positive

FN is False Negative

Sensitivity: The probability of correctly detecting that the subject suffers from the disease. A higher sensitivity means that the predictive model can easily detect the disease.

Sensitivity = TP/ (TP + FN) * 100 %

Specificity: Compared to sensitivity, specificity represents the probability of correctly determining that the subject does not suffer from the disease.

Specificity = TN/ (FP+ TN) * 100 %

Table 1: Performance Analysis for Various Patients

Metrics	Patient A	Patient B	Patient C
Accuracy	94.65	96.06	96.51
Sensitivity	95.75	96.42	96.06
Specificity	92.62	95.13	97.23

The Figure 2 shows the performance of the proposed methodology in terms of accuracy, sensitivity and specificity. From the resultant data, it is found that, the proposed methodology gives maximum of 96.51% accuracy, 96.42 % of sensitivity and 97.23% of specificity. It is concluded that the proposed methodology efficiently predict the diseases with respect the symptoms of the patients.

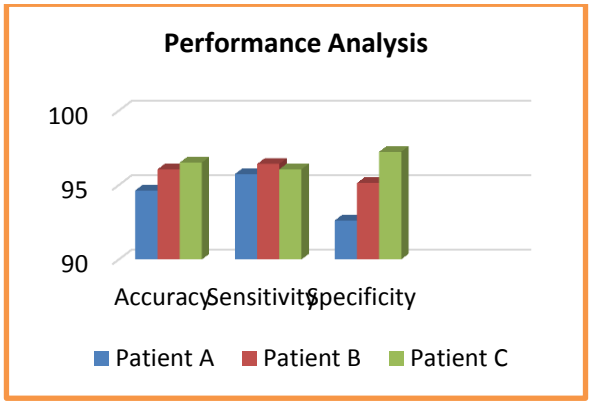

Figure 2: Performance Analysis for various Patient

The performance metrics are evaluated in two different data tests. Figure 3 shows the result of performance metrics of two data sets. From the figure 3, it is concluded that, the proposed methodology provides the maximum accuracy of 98.67 in PTR measure.

PTR: It is the proportion of positive test results that are true positives such as correct diagnoses. It is a critical measure of the performance of a diagnostic method, as it reflects the probability that a positive test reflects the underlying condition being tested for.

PTR = TP/ (TP+FP***) * 100 %***

NTR: It describes the performance of a diagnostic testing procedure. It is defined as the proportion of subjects with a negative test result who are correctly diagnosed. A high NPV for a given test means that when the test yields a negative result, it is most likely correct in its assessment.

NTR= TN/(TN+FN) * 100 %

The true positive rate represents the fraction of positive cases that are correctly classified by the model. The false positive rate represents the fraction of negative cases that are incorrectly classified as positive. Therefore, it provides a trade-off between sensitivity and specificity.

Table 2: Performance comparison

Performance Index	Dataset 1	Dataset 2
Accuracy	96.06	97.73
Sensitivity	95.52	97.32
Specificity	96.81	98.48
Positive Test Result	97.75	98.67
Negative Test Result	92.88	96.04

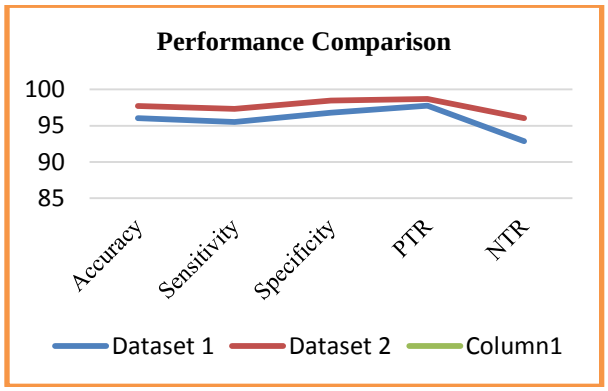

Figure 3: Performance Comparison

V. CONCLUSION

Health care plays a major role with the benefit of information technology. The disease diagnosis using mining is challenging due to its high risk, massive information, irrelevant and noisy data. The proposed automatic disease prediction system has explore knowledge using big data techniques and provides great potential to overcome those issues and improves clinical prediction almost perfect. Here input symptoms are given which extracts useful and relevant information from the database through apriori and frequent mining technology. Along with the symptoms by applying naive bayes algorithm, patient's history from EHR has been evaluated to make our results better. Although this model could accurately predict some diseases, it is still hard for some other diseases, especially those faces complexity in diagnosis. Hence, it urges them to collects the previous health records, doctor's prescription, and lab test values to improve the performance of the patient and produces qualitatively outcomes and provides accurate and efficient results, which will help patients get diagnosis instantly. The system reduces the human effort, cost and time constraint and increase the diagnostic accuracy. The overall mission of system development is to make the primary treatment quickly and easily with the use of technology. As it is said prevention is better than cure so the system will help the patient to let them know what they are suffering from till the doctor reacts to it. In future paramedical recommendation and online consultation with the doctors can be made for more convenience.

REFERENCES

[1]. Min Chen, Yixue Hao, Kai Hwang, Lu Wang, and Lin Wang, "Disease Prediction by Machine Learning over Big Data from Healthcare Communities", IEEE. 2169-3536, 2017.

[2]. Feixiang Huang, Shengyong Wang, and Chien-Chung Chan, "Predicting Disease By Using Data Mining Basedon Healthcare Information System", IEEE International Conference on Granular Computing, 978-1-4673-2311-6, 2012.

[3]. SujathaR ,Sumathy R and Anitha Nithya R, "A Survey of Health Care Prediction Using Data Mining", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, Issue 8, 2016.

[4]. Pinky Saikia Dutta, Shrabani Medhi, Sunayana Dutta, Tridisha Das and Sweety Buragohain, "Smart Health Care Using Data Mining", International Journal of Current Engineering And Scientific Research, ISSN : 2393-8374, Vol.-4, Issue-8,2017.

[5]. Ravi Aavula, M.Kruthini, N.Raviteja and K.Shashank, "Smart Health Consulting Android System", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Issue 3, 2017.

[6]. Priyanka Vijay Pawar, MeghaSakharamWalunj and PallaviChitte, "Estimation based on Data Mining Approach for Health Analysis", International Journal on Recent and Innovation Trends in Computing and Communication, Volume: 4 , 2017.

[7]. V. Krishnaiah, G. Narsimha and N. SubhashChandra, "A Study On Clinical Prediction Using Data Mining Techniques", International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR), Vol. 3, Issue 1, 2013.

[8]. Rawan Ali AL-jraib, Wafamoqhas, Huda Saeed Salman, AmeeraMohyAdden and Mona Mohmmad, "Biohouse Journal of Computer Science", International Journal Series, ISSN 2379-1500, 2017.

[9]. Md. Tahmid Rahman Laskar, Md. Tahmid Hossain, Abu RaihanMostofa Kamal and NafiulRashid, "Automated Disease Prediction System (ADPS): A User Input-based Reliable Architecture for Disease Prediction", International Journal of Computer Applications (0975 – 8887), Vol. 133 – No.15, 2016.

[10]. Aditya Tomar, "An Approach to Devise an Interactive Software Solution for Smart Health Prediction using Data Mining", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 7, 2016.

[11]. VidyaZope, PoojaGhatge, Aaron Cherian, PiyushMantri and KartikJadhav, "Smart Health Prediction using Machine Learning", IJSRD - International Journal for Scientific Research & Development| Vol. 4, Issue 12, 2017.

[12]. AakashKhatavkar, PiyushPotpose and PankajkumarPandey, "Smart Health Prediction System", IJSRD - International Journal for Scientific Research & Development Vol. 5, Issue 02, 2017.

[13]. Prashant Tiwari, AmanJaiswal, NarendraVishwakarm and PushpanjaliPatel, "Smart Health Care - An Android App To Predict Disease On The Basis Of Symptoms", International Research Journal of Engineering and Technology (IRJET), Vol.: 04 Issue: 04, 2017.

[14]. EvaK.Lee and Tsung-LinWu, "Classification and disease prediction via mathematical programming", American Institute of Physics, AIP Conference Proceedings, doi: 10.1063/1.2817343, 2007.

[15]. Riccardo Miotto, Li Li, Brian A. Kidd and Joel T. Dudley, "Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records", DOI: 10.1038 / srep26094, 2016.

[16]. Prasan Kumar Sahoo, Suvendu Kumar Mohapatra and Shih-Lin Wu, "Analyzing Healthcare Big Data with Prediction for Future Health Condition", IEEE OI 10.1109/ACCESS.2016.2647619, 2016.

[17]. Cheng-HsiungWeng, Tony Cheng-Kui Huang and Ruo-Ping Han, "Disease prediction with different types of neural network Classifiers", Elsevier Ltd. http://dx.doi.org/10.1016/j.tele.2015.08.006, 2016.

[18]. Xianglin Yang, Yunhai Ton, XiangfengMeng, Shuai Zhao, ZhiXu, YanjunLi,Guozhen Liu and Shaohua Tan, "Online Adaptive Method for Disease Prediction Based on Big Data of Clinical Laboratory Test", IEEE978-1-4673-9904-3, 2016.

[19]. AjinkyaKunjir, HarshalSawant, NuzhatF.Shaikh, "Data Mining and Visualization for Prediction of Multiple Diseases in Health Care", IEEE 978-1-5090-6399-4, 2017.

Authors Profile

S Manimekalai has completed her doctorate degree under Mother Teresa Women's University, Kodaikannal, Tamil Nadu, India. She has started her career as System Analyst (2002), in HUDCO, Chennai. Currently she is working as Assistant Professor in Theivanai Ammal College for Women (Autonomous), Villupuram. She has published two books and 19 National/ International Journals and presented papers in various conferences. Having 15 years of teaching experience. She has received Republic Day Achievers Award - 2018 as "Best Faculty Computer Science Software Reengineering" for year 2017. Her area of interest includes Digital Image Processing, Software Reengineering, Bigdata Analytics.

R.Suguna has completed her doctorate degree under Anna University, Chennai, Tamil Nadu, India. Currently she is working as Assistant Professor in Theivanai Ammal College for Women (Autonomous), Villupuram. She has published many National and International Journals and presented papers in various conferences. She has 10 years of teaching experience. Her area of interest includes Data Mining, Software Engineering and Networking.

S.Arulselvarani has completed her doctorate degree under Mother Teresa Women's University, Kodaikannal, Tamilnadu, India(2014). Currently she is working as Assistant Professor in STET Women's College, Mannargudi. She has published many National and International Journals and presented papers in various conferences. She has 14 yearsof teaching experience. She is a member in IEEE. Her area of interest includes E-Learning, Simulation,Datamining and Machine Learning.

