March-April 2020
ISSN: 0193-4120 Page No. 3350 - 3354

Similarity Disease Prediction System for Efficient Medicare

K.Mythili, S.Muthulakshmi, * T.Rajesh Kumar, T.Sangeetha

[1,2,4]Assitant Professor, Department of Information Technology, Sri Krishna College of Technology, Coimbatore, India ,

[3]Associate Professor, Department of Information Technology, Sri Krishna College of Technology, Coimbatore, India *rajeshkumarprofessor@gmail.com

Article Info
Volume 83
Page Number: 3350 - 3354
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 22 March 2020*

Abstract:
Disease prediction plays a key research area in the area of data mining. Predicting the serious effects of a disease at an early stage can decrease the death ratio and the healthcare industries focus more on this aspect. The technological development has now paved a great way for this. The major objective of the paper is to use the data mining techniques and the intelligent system efficiently for the prediction of patient's disease. The paper effectively does the prediction by comparing the already existing patient's with same symptoms so that providing medical solutions is also easy for the doctors. Analyzing EHR (Electronic Health Record) is a tedious process because it has a huge amount of data with various records. So a novel approach is proposed in the paper which will analyze and group the similar patients.

Keywords: *EHR, Intelligent System, Similar Patients.*

I INTRODUCTION

Data mining is the method for finding unknown values from enormous amount of data. The increase in patient population leads to increase in medical datasets. The transactions and investigation of these medical data is difficult without the computer based analysis system. The computer based analysis system indicates the mechanized medical diagnosis system. The mechanized diagnosis system support the medical practitioner to make good decision in treatment and disease. Data mining is the huge platform for the doctors to handle the huge amount of patient's datasets in many ways such as make sense of complex diagnostic tests, interpreting previous results, and combining the dissimilar data together.

II LITERATURE SURVEY

In the paper **"Disease Prediction System using data mining techniques"[1]** the author has discussed about the data mining techniques like association rule mining, classification, clustering to analyze the different kinds of heart based problems. The database used contain collection of records, each with a single class label, a classifier performs a brief and clear definition for each class that can be used to classify successive records. The data classification is based on MAFIA algorithms which result in accuracy, the data is estimated using entropy based cross validations and partition techniques and the results are compared C4.5 algorithm is used as the training algorithm to show rank of heart attack with the decision tree. The heart disease database is clustered using the K-means clustering algorithm, which will remove the data applicable to heart attack from the database. Some limitations are faced by the system like, time complexity is more due to DFS traversal, C4.5-Time complexity increases while searching for insignificant branches and lastly no precautions are defined.

In the paper "A study on data mining prediction techniques in healthcare sector"[2**]** the fields which discussed are Knowledge Discovery Process (KDD) is the process of changing the low-level data into

March-April 2020
ISSN: 0193-4120 Page No. 3350 - 3354

high-level knowledge. Hence, KDD refers to the nontrivial removal of implicit, previously unknown and potentially useful information from data in databases. The Knowledge Discovery in Databases process comprise of a few steps leading from raw data collections to some form of new information. The iterative process consists of the following steps: Data cleaning, Data integration, Data selection, Data transformation, Data mining, Pattern evaluation, Knowledge. Healthcare data mining prediction based on data mining techniques are as follows: Neural network, Bayesian Classifiers, Decision tree, Support Vector Machine. The paper states the comparative study of different healthcare predictions, Study of data mining techniques and tools for prediction of heart disease, various cancers, diabetes, eye disease and dermatological conditions. Data mining based prediction system reduces the human effects and cost effective one. Few limitations are that if attributes are not related then Decision trees prediction is less accurate and ANN is computationally intensive to train also it does not leads to specific conclusion.

The paper "An approach to devise an Interactive software solution for smart health prediction using data mining" [5] aims in developing computerized system to check and maintain your health by knowing the symptoms. It has a symptom checker module which actually defines our body structure and gives us liability to select the affected area and checkout the symptoms. Technologies implemented in this paper are: The front end is designed with help of HTML, Java Script and CSS. The back end is designed using MySQL which is used to design the databases. This paper also contains the information of testing like Alpha testing which is done at server side or we can say at the developer's end, this is an actual testing done with potential users or as an independent testing process at server end. And Beta testing is done after performing alpha testing, versions of a system or software known as beta versions are given to a specific audience outside the programming team. Only the limitation of this paper is it suggests only the award winning doctors and not the nearby doctors to the patient.

III EXISTING SYSTEM

There are many health care solutions provided using machine learning and data mining algorithms. The different algorithms are used to analyze single disease. For E.g.: Prediction of heart disease uses different algorithm whereas prediction of lung disease will uses different system. So, the complexity of system increases and for each type of disease we have to go for relevant system and find a matching solution. Instead of it, if a single system can analyze multiple diseases, then the efficiency will be good. So the intelligent system is developed which will analyze the new patient and compare his attributes with EHR and find the similar patient. Once similar patient is identified then their medical history will be classified and Medicare solution will be given accordingly.

IV PROPOSED SYSTEM

The proposed system contains 4 modules

1) EHR which stores previous patient's medical history.

2) Data Collection of New Patient.

3) Comparative Analysis and Probability of occurrence of disease.

4) Suitable Medications to be provided for patients.

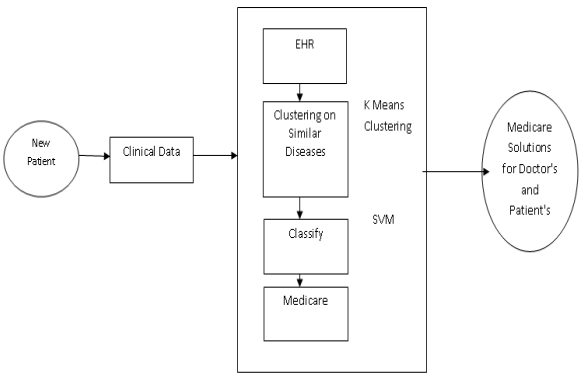

Figure 1. Architecture Diagram

March-April 2020
ISSN: 0193-4120 Page No. 3350 - 3354

The proposed system will not only predict the diseases but will also suggest the appropriate medications. The datasets will be used for both the symptoms checking, prediction of diseases and for providing medications and treatment. This system helps the doctor for the better treatment of patient's. This system will predict some major diseases like Cancer, liver and heart disease. This study aims mainly for the doctor and helps the doctor to treat the new patient based on the history of old patient record. This helps to provide suggestion for efficient Medicare.

1) EHR, which stores previous patient's medical history

Electronic Health record plays a vital role in storing all health information of various patients. It has all data like the personal and medical history, symptoms and the disease ,treatments provided for them, the improvements shows at each stage of treatment, final output of the treatment.

2) Data Collection of New Patient

The new patient's data are collected. The data includes age, disease, blood group, BP Level, historic information, Life style, working environment and so on. The attributes are compared with the existing EHR. The EHR contains discharge summary of various patients affected by the same or various disease. The system analyses and filters the matching disease and compares the attributes. The matching pattern is analyzed and clustered to groups.

3) Comparative Analysis and Probability of occurrence of disease

As the result of comparative analysis made, the patients having same symptoms are grouped together. With the clustered data the doctor's can understand the severity of disease and what type of medications to provide. The similarity

4) Suitable Medications to be provided for patients

Analyzing a new patient from the first level and testing each and every stage to conclude the level of disease then providing medications are complex and consume a lot of time. This time consuming process may sometime lead to increase in severity of disease. So a system is introduced which will predict the disease at early stage by analyzing the similar symptoms from other patients and already available EHR. The system helps the patient to get treated easily and provide solution for doctor's regarding Medicare. Many serious diseases if predicting at early stage can be treated and death rate can be reduced.

V EXPERIMENTAL RESULTS

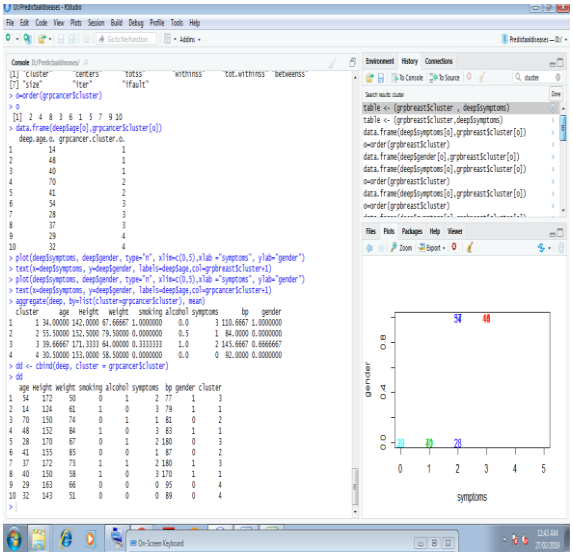

Figure 2. New Patient Clustering

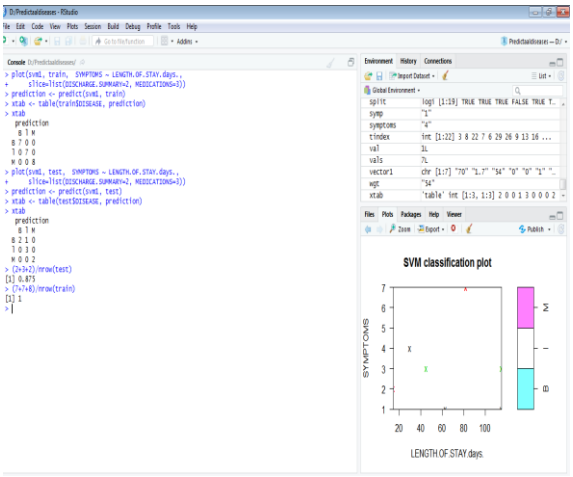

Figure 3. SVM Classification Plot

March-April 2020
ISSN: 0193-4120 Page No. 3350 - 3354

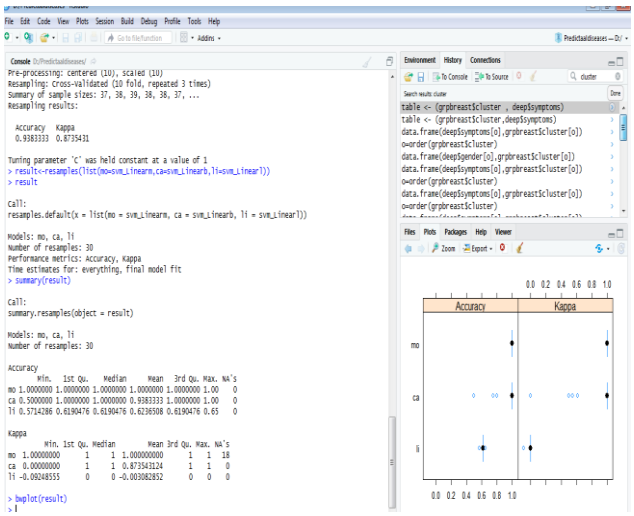

Figure 4. Accuracy Graph

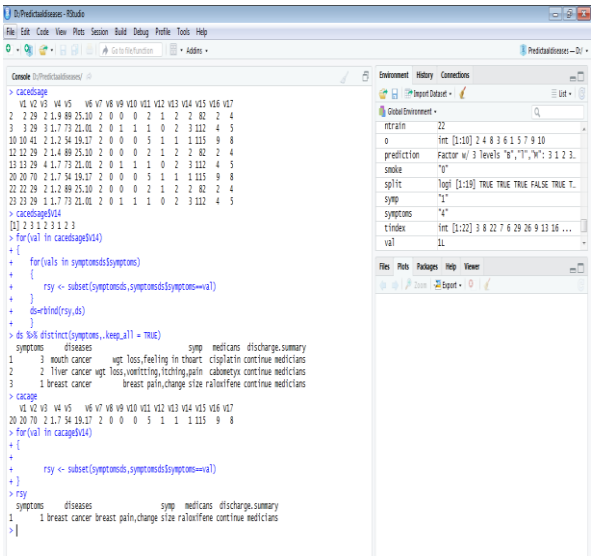

Figure 5. Disease Prediction

VI CONCLUSION & FUTURE ENHANCEMENT

The system thus provides a better solution for doctors and patients by easily identifying the disease and it's possible solution. The similarity prediction system is now efficient in analyzing the new patient record with EHR and finding the probability of similar disease .This helps to improve the Medicare solution to doctors. The time consumed in identifying the solution for diseases can be reduced. In future the system can be implemented by increasing the accuracy of the system and by considering some more diseases to provide medications[23].

VII REFERENCES

[1] K. Ng, J. Sun, J. Hu, and F. Wang, "Personalized predictive modeling and risk factor identification using patient similarity," *AMIA Summits onTranslational Science Proceedings*, 2015.

[2] N. Kasabov, "Global, local and personalised modeling and pattern discovery in bioinformatics: An integrated approach," *Pattern Recognition Letters*, 2007.

[3] J. Lee, D. M. Maslove, and J. A. Dubin, "Personalized mortality prediction driven by electronic medical data and a patient similarity metric," *PloS one'15*.

[4] J. Xu, J. Zhou, and P.-N. Tan, "Formula: Fac orized multi-task learning for task discovery in personalized medical models," in *SDM'15*.

[5] C. Che, C. Xiao, J. Liang, B. Jin, J. Zho, and F. Wang, "An rnn architecture with dynamic temporal matching for personalized predictions of parkinson's disease," in *SDM'17*. SIAM.

[6] Rajesh kumar,T., Geetha,K., "A Perspective Approach on Artificial Cognitive Computing and its Future Development", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 11, November 2016, ISSN(Online): 2320-9801, ISSN (Print): 2320-9798

[7] J. Sun, F. Wang, J. Hu, and S. Edabollahi, "Supervised patient similarity measure of heterogeneous patient records," *ACM SIGKDD Explorations Newsletter*, 2012.

[8] A. Gottlieb, G. Y. Stein, E. Ruppin, R. B. Altman, and R. Sharan, "Amethod for inferring medical diagnoses from patient similarities," *BMC medicine*, 2013.

[9] M. Zhan, S. Cao, B. Qian, S. Chang, and J. Wei, "Low-rank sparse feature selection for patient similarity learning," in *ICDM'16*.

[10] Q. Suo, H. Xue, J. Gao, and A. Zhang, "Risk factor analysis based on deep learning models," in *ACM-BCB'16*.

[11] Y. Cheng, F. Wang, P. Zhang, and J. Hu, "Risk prediction with electronic health records: A deep learning approach," in *SDM'16*.

[12] F. Ma, R. Chitta, J. Zhou, Q. You, T. Sun, and J. Gao, "Dipole: Diagnosis prediction in healthcare

March-April 2020
ISSN: 0193-4120 Page No. 3350 - 3354

via attention-based bidirectional recurrent neural networks," in *SIGKDD'17*.

[13] Selvy P.T., Palanisamy V., Purusothaman T. "Performance analysis of clustering algorithms in brain tumor detection of MR images2017. European Journal of Scientific Research 2011.

[14] R.Saranava Ram, M.Vinoth Kumar, S.Ramamoorthy, B.Saravana Balaji and Rajesh Kumar,T.,"An Efficient Hybrid Computing Environment to Develop a Confidential and Authenticated IoT Service Model" Wireless Personal Communications,doi.org/10.1007/s11277-020-07056-0, ISSN online: 1572-834X, Print ISSN : 0929-6212.

[15] Rajesh Kumar,T, Geetha, K, Remmiya Devi,G, Barkath Nisha,S, "Process Efficient Artificial Neural Network-Based Approach for Channel Selection and Classification of Seizures", Computational Network Application Tools for Performance Management, Asset Analytics, doi.org/10.1007/978-981-32-9585-8_12/

[16] Rajesh Kumar Thangavel.,Sveadha Athithan, Sarumathi, S, Aruna, M, Nithila, B, "Blackspot Alert and Accident Prevention System" 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE Xplore, 10.1109/ICCCNT45670.2019.8944412.

[17] Rajesh kumar,T Samsudeen, S, Sangeetha, S, Sudha Rani, " Enhanced Visual Attendance System by Face Recognition USING K-Nearest Neighbor Algorithm" Journal of Advanced Research in Dynamical & Control Systems, Vol. 11, 06-Special Issue, 2019

[18] Rajesh kumar,T., Remmiya devi,G., Abinaya,K., Deepika ,N N., Priyadharshini,S., "An Integrated Density Based Traffic Load Balancing System in a Cloud Environment", Pak. J. Biotechnol. Vol. 14 (4) 623-627 (2017), ISSN Print: 1812-1837, ISSN Online: 2312-7791

[19] Rajesh kumar, T., Preethi, S., Siva Rubini, R., Yamini, V., "SPEED DETECTING AND REPORTING SYSTEM USING GPS/GPRS AND GSM" International Journal of Pure and Applied Mathematics, Volume 118 No. 20 2018, 73-79, ISSN Online: 1314-3395

[20] Y. Sha, J. Venugopalan, and M. D. Wang, "A novel temporal similarity measure for patients based on irregularly measured data in electronic health records," *Platelets*, 2016.

[21] A. Sharafoddini, J. A. Dubin, and J. Lee, "Patient similarity in prediction models based on health data: a scoping review," *JMIR medical informatics*, 2017.

[22] Rajesh kumar, T.,Janani, A, Karthika, S, Raghunath, R, Samsudeen, U, Sai Vishnu, S, Krishna Rikachand Napolean," Automatic Seizure Detection in EEG signal using PSO and SVM "International Journal of Pure and Applied Mathematics, Volume 118 No. 20 2018, 239-245, ISSN Online: 1314-3395.

[23] Radha , B.Meena Preethi "Machine Learning Approaches For Disease Prediction From Radiology And Pathology Reports"", Journal of Green Engineering, Alpha Publishers, Vol.9,No.2,pp.149-166,2019

[24] Monterosso, D.M., Kumar, V. and Zala, K., 2019. Spiritual Practices in The Era of Smartphones & Social Networking: A Comparative Study. International Journal of Psychosocial Rehabilitation. Vol 22 (2) 45, 57.

[25] Shafti, S.S. and Ahmadie, M., 2018. Improvement of Psychiatric Symptoms by Cardiac Rehabilitation in Coronary Heart Disease Vol 22 (2) 80, 89.

[26] Ritter, V.C., Nordli, H., Fekete, O.R. and Bonsaksen, T., 2017. User satisfaction and its associated factors among members of a Norwegian clubhouse for persons with mental illness. International Journal of Psychosocial Rehabilitation. Vol 22 (1) 5, 14.

[27] Ferrazzi, P., 2018. From the Discipline of Law, a Frontier for Psychiatric Rehabilitation. International Journal of Psychosocial Rehabilitation, Vol 22(1) 16, 28.

[28] Bornmann, B.A. and Jagatic, G., 2018. Transforming Group Treatment in Acute Psychiatry: The CPA Model. International Journal of Psychosocial Rehabilitation, Vol 22(1) 29, 45.