IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

HEARTDISEASE PREDICTION USINGBIO INSPIRED ALGORITHMS

[1]P MOUNIKA, [2]V. VINAY KUMAR

[1](Assistant Professor), MSC, DNR college(A) PG courses Bhimavaram

[2]MSC, scholar, DNR college(A) PG courses Bhimavaram

ABSTRACT

Heart related diseases or Cardiovascular Diseases (CVDs) are the main reason for a huge number of death in the world over the last few decades and has emerged as the most lithreatening

disease, not only in India but in the whole world. So, there is a need of reliable, accurate afeasible system to diagnose such diseases in time for proper treatment. Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the health care industry and the professionals in the diagnosis of heart related diseases. This paper presents a survey of various models based on such algorithms and techniques and analyze their performance. Models based on supervised learning algorithms such as Support Vector Machines (SVM), K-Nearest Neighbor (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very popular among the researchers.

1.INTRODUCTION

According to a report by McKinsey [1], 50% of Americans have one or more chronic diseases, and 80% of American medical care fee is spent on chronic disease treatment. With the improvement of living standards, the incidence of chronic disease is increasing. The United States has spent an average of 2.7 trillion USD annually on chronic disease treatment. This amount comprises 18% of the entire annual GDP of the United States. The healthcare problem of chronic diseases is also very important in many other countries. In China, chronic diseases are the main cause of death, according to a Chinese report on nutrition and chronic diseases in 2015, 86.6% of

ISSN 2347–3657

Volume 12, Issue 3, 2024

deaths are caused by chronic diseases. Therefore, it is essential to perform risk assessments for chronic dis- eases. With the growth in medical data [2], collecting electronic health records (EHR) is increasingly convenient [3]. Besides, [4] rst presented a bio-inspired high-performance heterogeneous vehicular telematics paradigm, such that the collection of mobile users' health-related real-time big data can be achieved with the deployment of advanced hetero- geneous vehicular networks. Chen et al.proposed a healthcare system using smart clothing for sustainable health monitoring. Qiu et al. [8] had thoroughly studied the het- erogeneous systems and achieved the best results for cost minimization on tree and simple path cases for heteroge- neous systems. Patients' statistical information, test results and disease history are recorded in the EHR, enabling us to identify potential data-centric solutions to reduce the costs of medical case studies.Wang et al. [9] proposed an efcient ow estimating algorithm for the telehealth cloud system and designed a data coherence protocol for the PHR(Personal Health Record)-based distributed system. Bates et al. [10] proposed six applications of big data in the eld of health- care. Qiu et al. [11] proposed an optimal big data sharing algorithm to handle the complicate data set in telehealth with cloud techniques. One of the applications is to identify high-risk patients which can be utilized to reduce medical cost since high-risk patients often require expensive healthcare. Moreover, in therst paper proposing health- care cyber-physical system [12], it innovatively brought for- ward the concept of prediction-based healthcare applications, including health risk assessment. Prediction using traditional disease risk models usually involves a machine learning algorithm (e.g., logistic regression and regression analysis, etc.), and especially a supervised learning algorithm by the use of training data with labels to train the model [13], [14]. In the test set, patients can be classified into groups of either high-risk or low-risk. These models are valuable in clinical situations and are widely studied [15], [16]. However, these schemes have the following characteristics and defects. The data set is typically small, for patients and diseases with specific conditions [17], the characteristics are selected through experience. However, these pre-selected characteristics maybe not satisfy the changes in the disease and its influencing factors.

ISSN 2347–3657

Volume 12, Issue 3, 2024

With the development of big data analytics technology, more attention has been paid to disease prediction from the perspective of big data analysis, various researches have been conducted by selecting the characteristics automatically from a large number of data to improve the accuracy of risk clas- sication [18], [19], rather than the previously selected characteristics. However, those existing work mostly considered structured data. For unstructured data, for example, using convolutional neural network (CNN) to extract text characteristics automatically has already attracted wide attention and also achieved very good results [20], [21] . However, to the best of our knowledge, none of previous work handle Chinese medical text data by CNN. Furthermore, there is a large difference between diseases in different regions, primarily because of the diverse climate and living habits in the region. Thus, risk classification based on big data analysis, the fol- lowing challenges remain: How should the missing data be addressed? How should the main chronic diseases in a certain region and the main characteristics of the disease in the region be determined? How can big data analysis technology be used to analyze the disease and create a better model?

To solve these problems, we combine the structured and unstructured data in healthcare eld to assess the risk of disease. First, we used latent factor model to reconstruct the missing data from the medical records collected from a hospital in central China. Second, by using statistical knowledge, we could determine the major chronic diseases in the region. Third, to handle structured data, we consult with hospital experts to extract useful features. For unstruc- tured text data, we select the features automatically using CNN algorithm. Finally, we propose a novel CNN-based multimodal disease risk prediction (CNN-MDRP) algorithm for structured and unstructured data. The disease risk model is obtained by the combination of structured and unstructured features. Through the experiment, we draw a conclusion that the performance of CNN-MDPR is better than other existing methods.

2.EXISTING SYSTEM

Heart diseases have emerged as one of the most prominent cause of death all around the world. According to World Health Organization, heart related diseases

ISSN 2347–3657

Volume 12, Issue 3, 2024

are responsible for the taking 17.7 million lives every year, 31% of all global deaths. In India too, heart related diseases have become the leading cause of mortality [1]. Heart diseases have killed 1.7 million Indians in 2016, according to the 2016 Global Burden of Disease Report, released on September 15,2017. Heart related diseases increase the spending on health care and also reduce the productivity of an individual. Estimates made by the World Health Organization (WHO), suggest that India have lost up to $237 billion, from 2005-2015, due to heart related or Cardiovascular diseases . Thus, feasible and accurate prediction of heart related diseases is very important.

Medical organizations, all around the world, collect data on various health related issues. These data can be exploited using various machine learning techniques to gain useful insights. But the data collected is very massive and, many a times, this data can be very noisy. These datasets, which are too overwhelming for human minds to comprehend, can be easily explored using various machine learning techniques. Thus,

these algorithms have become very useful, in recent times, to predict the presence or absence of heart related diseases accurately.

3.PROPOSED SYSTEM

Dimensionality Reduction involves selecting a mathematical representation such that one can relate the majority of, but not all, the variance within the given data, thereby including only most significant information. The data considered for a task or a problem, may consists of a lot of attributes or dimensions, but not all of these attributes may equally influence the output. A large number of attributes, or features, may affect the computational complexity and may even lead to overfitting which leads to poor results. Thus, Dimensionality Reduction is a very important step considered while building any model. Dimensionality Reduction is generally achieved by two methods -Feature Extraction and Feature Sel

4.SCREENSHOTS

Home Page

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 3, 2024

User Registration Page

Admin login

Admin Home Page

Admin View Registered Users

User Login

User Home Page

5.CONCLUSION

In this paper, we propose a new 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

ISSN 2347–3657

Volume 12, Issue 3, 2024

speed which is faster than that of the CNN-based unimodal disease risk prediction (CNN-UDRP) algorithm

6.REFERENCES

[1] P. Groves, B. Kayyali, D. Knott, and S. van Kuiken, TheBig DataRevolution in Healthcare: Accelerating Value and Innovation. USA: Center for US Health System Reform Business Technology Of ce, 2016.

[2] M. Chen, S. Mao, and Y. Liu, Big data: A survey, Mobile Netw. Appl., vol. 19, no. 2, pp. 171209, Apr. 2014.

[3] P. B. Jensen, L. J. Jensen, and S. Brunak, Mining electronic health records: Towards better research applications and clinical care, Nature Rev. Genet., vol. 13, no. 6, pp. 395405, 2012.

[4] D. Tian, J. Zhou, Y. Wang, Y. Lu, H. Xia, and Z. Yi, A dynamic and self-adaptive network selection method for multimode communications in heterogeneous vehicular telematics, IEEE Trans. Intell. Transp. Syst., vol. 16, no. 6, pp. 30333049, Dec. 2015.

[5] M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, and C. Youn, Wearable 2.0: Enable human-cloud integration in next generation healthcare system, IEEE Commun., vol. 55, no. 1, pp. 5461, Jan. 2017.

[6] M. Chen, Y. Ma, J. Song, C. Lai, and B. Hu, Smart clothing: Con necting human with clouds and big data for sustainable health monitor ing, ACM/Springer Mobile Netw. Appl., vol. 21, no. 5, pp. 825845, 2016.

[7] M. Chen, P. Zhou, and G. Fortino, Emotion communi cation system, IEEE Access, vol. 5, pp. 326337, 2017, doi: 10.1109/ACCESS.2016.2641480.

[8] M. Qiu and E. H.-M. Sha, Cost minimization while satisfying hard/soft timing constraints for heterogeneous embedded systems, ACM Trans. Design Autom. Electron. Syst., vol. 14, no. 2, p. 25, 2009.