International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

DESIGN AND IMPLEMENTATION OF MACHINE LEARNING BASED EARLY PROGNOSTICATION OF CHRONIC KIDNEY DISEASE

Nivetha R[1], Prof. Dr. Azha.Periyasamy[2]

[1]Nivetha Department of Electronics and Instrumentation & Bharathiar University

[2] Prof.Dr.Azha.Periasmay Department of Electronics and Instrumentation & Bharathiar University

---

Abstract - - End stage renal disease (ESRD) describes the most severe stage of chronic kidney disease (CKD), when patients need dialysis or renal transplant. There is often a delay in recognizing, diagnosing, and treating the various stages of CKD. The objective of the present study was to employ machine learning algorithms to develop a prediction model for progression to ESRD based on a large-scale multidimensional database. Machine learning (ML) techniques are excellent in predicting CKD. The current study offers a methodology for predicting CKD status using clinical data, which incorporates data preprocessing, data balancing and model design. A technique for managing missing values is done with the help of data aggregation and feature extraction. A number of physiological variables, as well as ML techniques such as logistic regression (LR), decision tree (DT) classification, and k-nearest neighbor (KNN), were used in this work to train three distinct models for reliable prediction. Compared to prior research, the accuracy rate of the models employed in this study is considerably greater, implying that they are more trustworthy than the models used in previous studies as well. A large number of model comparisons have shown their resilience, and the scheme may be inferred from the study's result

Key Words- CKD, Machine learning, LR, DT,KNN.

1.INTRODUCTION

End stage renal disease (ESRD) describes the most severe stage of chronic kidney disease (CKD), when patients need dialysis or renal transplant. There is often a delay in recognizing, diagnosing, and treating the various stages of CKD. The objective of the present study was to employ machine learning algorithms to develop a prediction model for progression to ESRD based on a large-scale multidimensional database. Machine learning (ML) techniques are excellent in predicting CKD. The current study offers a methodology for predicting CKD status using clinical data, which incorporates data preprocessing, data balancing and model design. A technique for managing missing values is done with the help of data aggregation and feature extraction. A number of physiological variables, as well as ML techniques such as logistic regression (LR), decision tree (DT) classification, and k-nearest neighbor (KNN), were used in this work to train three distinct models for reliable prediction. Compared to prior research, the accuracy rate of the models employed in this study is considerably greater, implying that they are more trustworthy than the models used in previous studies as well. A large number of model comparisons have shown their resilience, and the scheme may be inferred from the study's result.

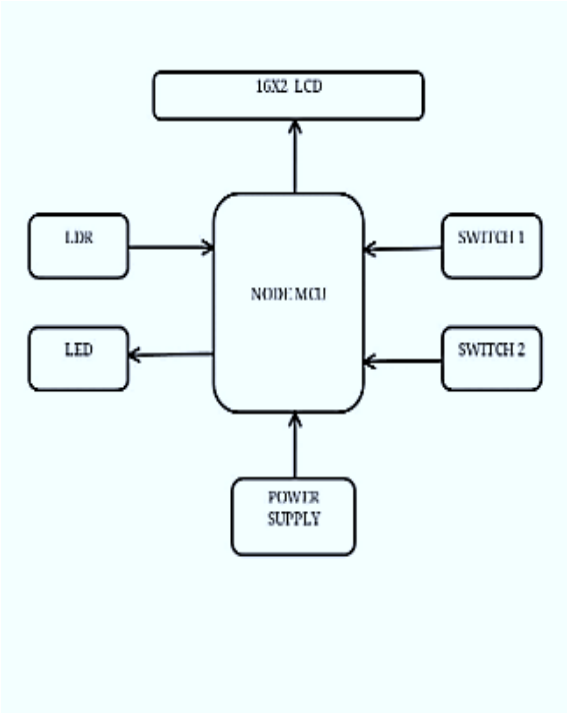
LITERATURE SURVEY

Chronic kidney disease (CKD) represents a heavy burden on the healthcare system because of the increasing number of patients, high risk of progression to end-stage renal disease, and poor prognosis of morbidity and mortality. The aim of this study is to develop a machine-learning model that uses the co morbidity and medication data obtained from Taiwan's National Health Insurance Research Database to forecast the 11 occurrence of CKD within the next 6 or 12 months before its onset, and hence its prevalence in the population. A total of 18,000 people with CKD and 72,000 people without CKD diagnosis were selected using propensity score matching. Their demographic, medication and comorbidity data from their respective two-year observation period were used to build a predictive model. Among the approaches investigated, the Convolution Neural Networks (CNN) model performed best with a test set AUROC of 0.957 and 0.954 for the 6-month and 12-month predictions, respectively. The most prominent predictors in the tree-based models were identified, including diabetes mellitus, age, gout, and medications such as sulfonamides and angiogenesis. The model proposed in this study could be a useful tool for policymakers in predicting the trends of CKD in the population. The models can allow close monitoring of people at risk, early detection of CKD, better allocation of resources, and patient-centric management[1]. CKD is a serious reason of demise and disability. It was the 27th focal reason in 1990 and became 18th focal reason in 2010. Near about 1 million people lose their life in 2013. In spite of that, people of developing countries are being affected by CKD. We analyzed the data of CKD patient and proposed a system from which it will be possible to predict the risk of CKD. We have used 455 patients' data. Online data set which is collected from UCI Machine Learning Repository and real time dataset which is collected from Khulna City Medical College are used here. We used Python as a high-level interpreted programming language for developing our system. We trained the data using 10-fold CV and applied Random forest and ANN. The accuracy achieved by Random forest algorithm is 97.12% and ANN is 94.5%. This system will help to predict early disclosure of chronic kidney diseases[2].

METHODOLOGY

This study reveals the results in three phases, i.e., preprocessing, computation and final results to predict the stages of chronic kidney disease. Block diagram of the proposed method is designed in software by the authors. The methods were devised in accordance with relevant guidelines and regulations. Preprocessing this phase starts from the acquisition of dataset of CKD patients. Four attributes are selected from the dataset to be given as input in GFR calculation. As, this equation is reliable for the calculation of all stages of CKD as compared to Modification of Diet in Renal

BLOCK DIAGRAM



HARDWARE COMPONENTS

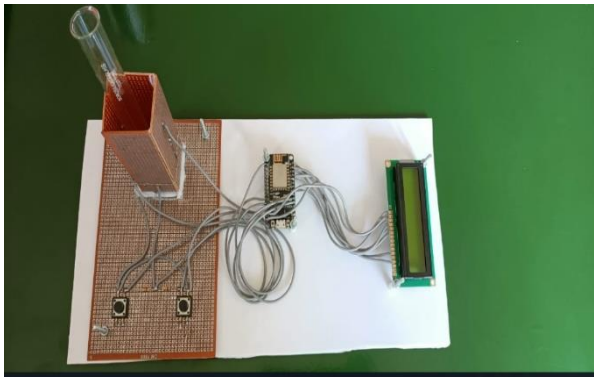
¬ NODEMCU (ESP8266)
¬ POWER SUPPLY
¬ 16X2 LCD
¬ LDR
¬ LED
¬ SWITCH

ARDUINO

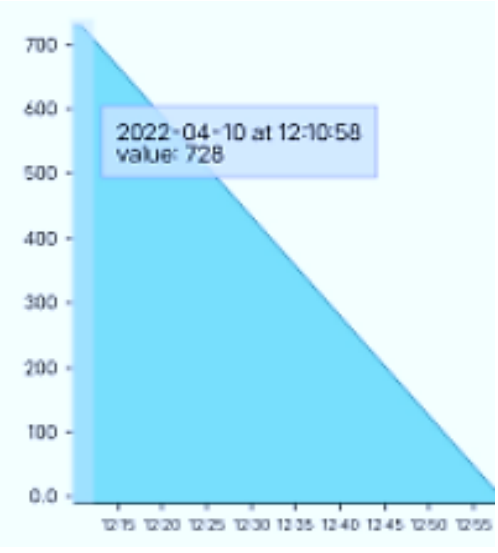
The open-source Arduino Software (IDE) makes it easy to write code and upload it to the board. This software can be used with any Arduino board. The Arduino Integrated Development Environment - or Arduino Software (IDE) - contains a text editor for writing code, a message area, a text console, a toolbar with buttons for common functions and a series of menus. It connects to the Arduino hardware to upload programs and communicate with them. Writing Sketches Programs written using Arduino Software (IDE) are called sketches. These sketches are written in the text editor and are saved with the file extension. The editor has features for cutting/pasting and for searching/replacing text. The message area gives feedback while saving and exporting and also displays errors. The console displays text output by the Arduino Software (IDE), including complete error messages and other information. The bottom right-hand corner of the window displays the configured board.

The ESP8266 is a low-cost Wi-Fi microchip, with built-in TCP/IP networking software, and microcontroller capability, produced by Expressive Systems in Shanghai, China. The chip was popularized in the English-speaking maker community in August 2014 via the ESP-01 module, made by a third-party manufacturer Ai-Thinker. This small module allows microcontrollers to connect to a Wi-Fi network and make simple TCP/IP connections using Hayes-style commands. However, at first, there was almost no English-language documentation on the chip and the commands it accepted. The very low price and the fact that there were very few external components on the module, which suggested that it could eventually be very inexpensive in volume, attracted many hackers to explore the module, the chip, and the software on it, as well as to translate the Chinese documentation. The ESP8285 is a similar chip with a built-in 1 MiB flash memory, allowing the design of single-chip devices capable of connecting via Wi-Fi. FEATURES Processor : L106 32-bit RISC microprocessor core based on the Tensilica Diamond Standard 106Micro running at 80 MHz

HARDWARE SETUP



RESULT



OUTPUT

Date	Value
2022-04-10 at 12:09:55	728
2022-04-10 at 12:09:51	729
2022-04-10 at 12:09:47	728
2022-04-10 at 12:09:43	728
2022-04-10 at 12:09:39	729
2022-04-10 at 12:09:35	730
2022-04-10 at 12:09:31	729
2022-04-10 at 12:09:27	729

FUTURE SCOPE

Chronic Kidney disease is a generic term that covers various heterogeneous kidney disorders. Five to ten percent of the population worldwide suffers from this disease. Chronic Kidney Disease is a worldwide health crisis. A majority of the cases of Chronic Kidney Disease go undiagnosed or are diagnosed later in underdeveloped and developing nations; this is one of the prime reasons that higher percentage of these cases are from developing and underdeveloped Nations as compared to developed nations where majority of people go through routine check-up and diagnosis. According to a report: “More than 80% of all patients who receive treatment for kidney failure are in affluent countries with universal access to health care & large elderly populations” In the future course of this study one can try to further improve the two-class classification accuracy by evaluating some hybrid or ensemble techniques, in addition to this a subset of features can be extracted from the complete medical data-set of chronic kidney disease of twenty four parameters (features) without effecting the performance of the classification process, so that the financial burden a patient has to bear for undergoing various clinical tests can be reduced.

CONCLUSION

We have analyzed 14 different attributes related to CKD patients and predicted accuracy for different machine learning algorithms like Decision tree and Support Vector Machine. From the results analysis, it is observed that the decision tree algorithms give the accuracy of 91.75% and SVM gives accuracy of 96.75%. When considering the decision tree algorithm it builds the tree based on the entire dataset by using all the features of the dataset. The advantage of this system is that, the prediction process is less time consuming. It will help the doctors to start the treatments early for the CKD patients and also it will help to diagnose more patients within a less time period. Limitations of this study are the strength of the data is not higher because of the size of the data set and the missing attribute values. To build a machine learning model targeting chronic kidney disease with overall accuracy of 99.99%, will need millions of records with zero missing value.

REFERENCE

1. Setu Basak, Md. Mahbub Alam, Aniruddha Rakshit, Ahmed Al Marouf, Anup Majumder Predicting and Staging Chronic Kidney Disease of Diabetes (Type-2) Patient Using Machine Learning Algorithms DOI: 10.35940/ijeat.A2213.109119\

2. Vinod Sharma, Atul Sharma ,Sahil Sharma Performance Based Evaluation of Various Machine Learning Classification Techniques for Chronic Kidney Disease Diagnosis

3. Mohamed Ahmed Alloghani,Dhiya Al-Jumeily Obi Performance-Based Prediction of Chronic Kidney Disease Using Machine Learning for High-Risk Cardiovascular Disease Patients DOI:10.1007/978-3-030-28553-1_9

4. Surya Krishnamurthy , Kapeleshh KS , Erik Dovgan , Mitja Luštrek , Barbara Gradišek Pileti˘c , Kathiravan Srinivasan , Yu-Chuan (Jack) Li , Anton Gradišek and Shabbir Syed-Abdul Machine Learning Prediction Models for Chronic Kidney Disease Using National Health Insurance Claim Data Taiwan https://doi.org/ 10.3390/healthcare 9050546

5. S.Ramya, Dr. N.Radha Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms DOI: 10.15680/IJIRCCE.2016. 0401049

6. S.Revathy, B.Bharathi, P.Jeyanthi, M.Ramesh Chronic Kidney Disease Prediction using Machine Learning Models DOI: 10.35940/ijeat.A2213.109119

7. Siddheshwar Tekale, Pranjal Shingavi, Sukanya Wandhekar, Ankit Chatorikar Prediction of Chronic Kidney Disease Using Machine Learning Algorithm DOI 10.17148/IJARCCE.2018.710213 37

8. Minhaz Uddin Emon, Al Mahmud Imran, Rakibul Islam, Maria Sultana Keya, Raihana Zannat Performance Analysis of Chronic Kidney Disease through Machine Learning ApproachesPDF DOI: 10.1109/ICICT50816.2021.9358491

9. Gazi Mohammed Ifraz , Muhammad Hasnath Rashid ,Tahia Tazin ,Sami Bourouis , and Mohammad Monirujjaman Khan Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods .

10. Hamida Ilyas1, Sajid Ali1, Mahvish Ponum1, Osman Hasan1 , Muhammad Tahir Mahmood1, Mehwish Iftikhar1 Chronic kidney disease diagnosis using decision tree algorithms BMC Nephrology (2021) 22:273

11. Shanila Yunus Yashfi, Md Ashikul Islam, Pritilata, Nazmus Sakib, Tanzila Islam Risk Prediction Of Chronic

Kidney Disease Using different classification Algorithms DOI: 10.1109/ICCCNT49239.2020.9225548

12. Jiongming Qin, Lin Chen, Yuhua Liu, Chuanjun Liu, Changhao Feng, Bin Chen A Machine Learning Methodology for Diagnosing Chronic Kidney Disease DOI: 10.1109/ACCESS.2019.2963053

13. Thipwan Fufaung Tippawan Niyomwong Wandee Chokchueypattanakit Sathit Suwannawach Nitat Ninchawee Predictive analytics for chronic kidney disease using machine learning techniques DOI: 10.1109/MITICON.2016.8025242