

Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease

Ch V Ramana¹, J.Vinod Kumar², N.Sravani³, M.Virajitha⁴, V.Vivek⁵
^{1,2,3,4,5}Pragati Engineering College

Abstract - chronic kidney disease (CKD) is a type of chronic disease which means it happens slowly over a period of time and persists for a long time thereafter. It is deadly at its end stage and will only be cured by kidney replacement or regular dialysis which is an artificial filtering mechanism. It is important to identify CKD at the early stage so that necessary treatments can be provided to prevent or cure the disease. The main focus in this paper is on the classification techniques, that is, tree-based decision tree, random forest, and logistic regression has been analyzed. Different measure has been used for comparison between algorithms for the dataset collected from standard UCI repository. That over a period of time and persists for a long time thereafter. It is deadly at its end stage and will only be cured by kidney replacement or regular dialysis which is an artificial filtering mechanism. It is important to identify CKD at the early stage so that necessary treatments can be provided to prevent or cure the disease.

I. INTRODUCTION

Chronic Kidney Disease (CKD) is a critical health condition worldwide that is a major reason for malicious health outcomes, particularly in countries where income ranges from low-to-middle where millions die regularly due to lack of modest treatment. As per the stages in any chronic disease the fatality is related to the stage it had been without being cured. The high-risk factors of CKD are increasing frequency of diabetic patient, hypertension, heart disease, mellitus and family history of kidney failure. If CKD is left undetected and therefore untreated, it can lead to hypertension and in severe cases to kidney failure. WE procured a standard dataset from the UCI machine repository for chronic Kidney Disease. CKD if predicted early and accurately, can benefit patients in many ways. It increases the probability of a successful treatment while also adding years to the person's life. This paperwork aims to predict kidney disease by using some of the selected machine learning

algorithms and feature selection methods. The objective is to collect the combination of different feature and then have used it as input to the machine learning algorithms. The algorithms have been implemented on the basis of selected features and then we compare their performances.

II. HARDWARE SOFTWARE SPECIFICATION

A. REQUIREMENT ANALYSIS

The project involved analyzing the design of few applications so as to make the application more users friendly. To do so, it was really important to keep the navigation from one screen to the other well-ordered and at the same time reducing the amount of typing the user needs to do

B. SYSTEM SPECIFICATION

HARDWARE REQUIREMENTS

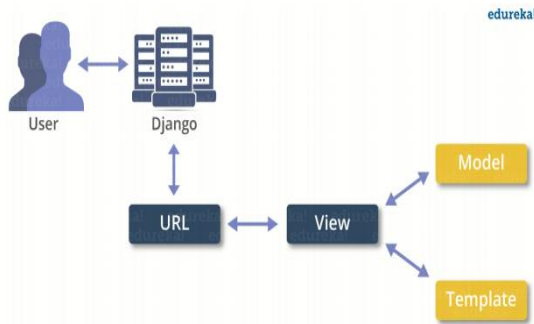
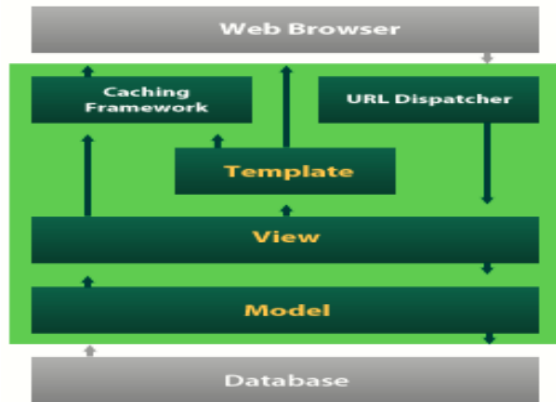
System : Pentium IV 2.4 GHz. ⌘ Hard Disk : 40 GB. ⌘ Floppy Drive : 1.44 Mb. ⌘ Monitor : 14' Colour Monitor. ⌘ Mouse : Optical Mouse. ⌘ Ram : 512 Mb. ⌘

SOFTWARE ENVIRONMENT

PYTHON Python is a general-purpose interpreted, interactive, objectoriented, and high-level programming language. An interpreted language, Python has a design philosophy that emphasizes code readability (notably using whitespace indentation to delimit code blocks rather than curly brackets or keywords), and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++ or Java.

DJANGO Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app

without needing to reinvent the wheel. It's free and open source.



Django Administration process

Prediction performance of individual and ensemble learn for chronic kidney disease

AUTHORS: Dili Singh Sisodia; Akanksha Verma

Automating the process of predicting diseases prove assistive and time-saving for a practitioner in the field of medical diagnosis. The accurate prediction of any disease not only helps the patients know about their health but also helps the doctors in medication suggestion well in advance. In today's lifestyle, advance knowledge about health and proper care can add a number of living days to a patient's life. In this paper, the prediction of chronic kidney disease (CKD) is performed using individual and ensemble learners. The experiments are performed on CKD dataset was taken from UCI repository. The three different classifiers from individual classifiers, namely, Naive Bayes(NB), minimal sequential optimization (SMO), J48, and three ensemble classifiers, namely, Random Forest (RF), bagging, AdaBoost respectively are used for prediction.

Chronic Kidney Disease analysis using data mining classification techniques:

Authors:Veenita Kunwar; Khushboo Chandel; A. Sai Sabitha; Abhay Bansal. Data mining has been a current trend for attaining diagnostic results. Huge amount of unmined data is collected by the healthcare industry in order to discover hidden information for effective diagnosis and decision making. Data mining is the process of extracting hidden information from massive dataset, categorizing valid and unique patterns in data. There are many data mining techniques like clustering, classification, association analysis, regression etc.

III. SYSTEM ANALYSIS

EXISTING SYSTEM: Cosmology and machine learning for Chronic Kidney Disease as a complex versatile WEKA tool. Ontology and machine learning are the techniques that have been utilized in existing methodology. Therefore, it shows a Chronic Kidney Disease to help instrument for taking care of mistakes in the and helps clinicians adequately recognize intense kidney torment patients from those with different reasons for kidney torments. Another machine learning procedure is Coronary Artery Disease method called N2 Genetic optimizer agent.

DISADVANTAGES OF EXISTING SYSTEM: Machine learning-based coronary artery disease examined datasets,→ test sizes, highlights, areas of information accumulation, execution measurements, and applied ML are the basic methods that have been broken down in this methodology. Chronic kidney Failure Detection is an anticipated the Constant kidney→ breakdown identification from heart sounds utilizing a pile of machine learning classifiers. The strategies used to foresee comprises filtering segmentation and feature extraction to the model. Algorithm: K-Nearest Neighbor, Naïve Baye

PROPOSED SYSTEM

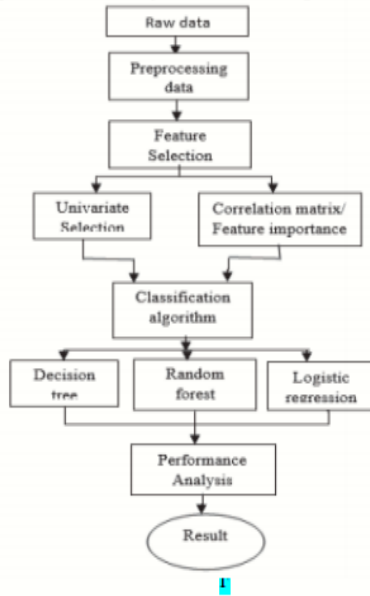
We worked on heart disease dataset obtained from UCI (University of California at Irvine) repository, the data set contained attributes such as age, sex, cp, Blood pressure, Blood urea, Sugar, Potassium, ca, and target with 401 instances has taken.

ADVANTAGES OF PROPOSED SYSTEM: The accuracy of the classifiers was calculated using the confusion matrix. The classifier which bags up the

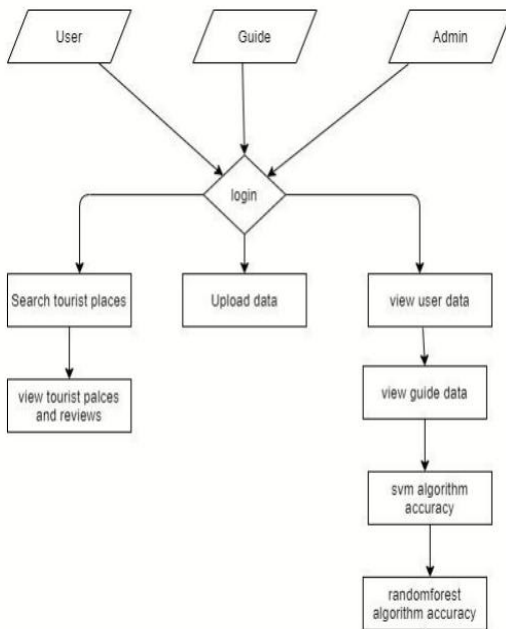
highest accuracy could be determined as the best classifier. Algorithm: LOGISTIC REGRESSION, Random Forest (RF), DECISSION TREE

IV. SYSTEM DESIGN

System architecture



Data Flow Diagram



SYSTEM STUDY

FEASIBILITY STUDY The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. For feasibility analysis, some understanding of the major requirements for the system is essential. Three key considerations involved in the feasibility analysis are, ECONOMICAL FEASIBILITY ♦ TECHNICAL FEASIBILITY ♦ SOCIAL FEASIBILITY ♦

ECONOMICAL FEASIBILITY This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited.

TECHNICAL FEASIBILITY This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources.

SYSTEM TEST

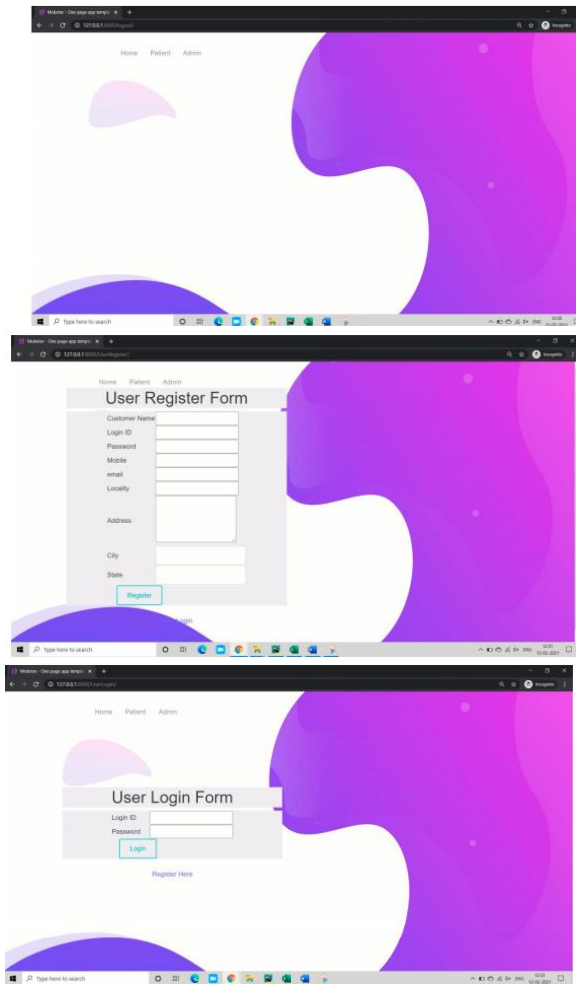
The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

RESULT AND TESTING:

The research study for Tourist Place review classification using machine learning algorithms has future scope of handling multilingual review classification. Also we will try to use different feature selection methods like Recursive feature elimination with cross-validation to improve accuracy of classification. In future work We will try to use it. deep learning based techniques for feature extraction and classification for better performance

OUTPUT SCREENSHORTS

HOMEPPAGE



User data

S.No	Name	Login ID	Mobile	Email	Locality	Status	Activate
1	arun	arun	9849843928	arun@gmail.com	hyd	activated	Activated
2	sagar	sagar	7262894008	sagar.datapoint21@gmail.com	hyd	activated	Activated

V.CONCLUSION

We were able to evaluate the performance of different ML algorithms on the Chronic Kidney Disease data set we took from UCI machine learning library [11]. We preprocess the dataset and then used the filter method of feature selection that is univariate selection and

correlation matrix along with feature importance to find best features from the dataset. The proposed algorithm that is, Decision tree, Random forest and logistic regression have achieved an accuracy of 98.48, 94.16 and 99.24 respectively. Precision of 100, 95.12 and 98.82 and recall of 97.61, 96.29 and 100. Two feature selecting techniques are combined by leveraging the strength of each the techniques. On comparison we find Logistic Regression with highest accuracy and recall while Decision tree have the highest precision

VI.ACKNOWLEDGMENT

I would like to express my gratitude to all the people behind the screen who helped me to transform an idea into a real application. I would like to express my heart-felt gratitude to my parents without whom I would not have been privileged to achieve and fulfill my dreams. I am grateful to our principal, Dr. S SAMBHU PRASAD who most ably run the institution and has had the major hand in enabling me to do my project. I profoundly thank Dr.D.SIRISHA, Head of the Department of Information Technology who has been an excellent guide and also a great source of inspiration to my work. I would like to thank my internal guide Mr.CHINTA VENKATA RAMANA for his technical guidance, constant encouragement and support in carrying out my project. The satisfaction and euphoria that accompany the successful completion of the task would be great but incomplete without the mention of the people who made it possible with their constant guidance and encouragement crowns all the efforts with success. In this context, I would like thank all the other staff members, both teaching and non-teaching, who have extended their timely help and eased my task. PROJECT ASSOCIATES J.VINOD KUMAR 17A31A1247 N.SRAVANI 17A31A1219 M.VIRAJITHA 17A31A1213 V.VIVEK 17A31A1257

REFERENCES

- [1] A. R, G. Sasi, R. Sankar, and O. Deepa, "Decision Support system for diagnosis and prediction of Chronic Renal Failure using Random Subspace Classification," 2016, pp. 1287–1292.
- [2] D. S. Sisodia and A. Verma, "Prediction Performance of Individual and Ensemble learners

- for chronic kidney disease,” 2017, pp. 1027–1031.
- [3] A. V Kshirsagar et al., “A Simple Algorithm to Predict Incident Kidney Disease,” *ARCH Intern Med*, vol. 168, no. 22, pp. 2466–2473, 2008.
- [4] A. K. Shrivastava and S. Kumar Sahu, “Classification of Chronic Kidney Disease using Feature Selection Techniques,” *IJCSE*, vol. 6, no. 5, pp. 649–653, 2018.
- [5] M. Kumar, “Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 5, no. 2, pp. 24–33, 2016.
- [6] P. Yildirim, “Chronic Kidney Disease Prediction on Imbalanced Data by Multilayer Perceptron: Chronic Kidney Disease Prediction,” in *Proceedings - International Computer Software and Applications Conference*, 2017, vol. 2, pp. 193–198, doi: 10.1109/COMPSAC.2017.84.
- [7] E. H. A. Rady and A. S. Anwar, “Prediction of kidney disease stages using data mining algorithms,” *Informatics Med. Unlocked*, vol. 15, pp. 1–7, Jan. 2019, doi: 10.1016/j.imu.2019.100178.
- [8] V. Kunwar, K. Chandel, S. A. Sai, and A. Bansal, “Chronic kidney disease analysis using data mining classification techniques,” 2016, pp. 300–305.
- [9] V. S and D. S, “Data Mining Classification Algorithms for Kidney Disease Prediction,” *Int. J. Cybern. Informatics*, vol. 4, no. 4, pp. 13–25, Aug. 2015, doi: 10.5121/ijci.2015.4402.
- [10] A. Nway Oo, “Classification of Chronic Kidney Disease (CKD) Using Rule based Classifier and PCA,” *Int. J. Adv. Manag. Technol. Eng. Sci.*, vol. 8, no. 4, pp. 728–733, 2018.
- [11] Dua, Dheeru and Graff, and Casey, “UCI Machine Learning Repository,” 2017. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease. [Accessed: 20- Mar-2020].
- [12] A. K. Shrivastava and S. Kumar Sahu, “Classification of Chronic Kidney Disease using Feature Selection Techniques,” *IJCSE*, vol. 6, no. 5, pp. 649–653, 2018.
- [13] M. Kumar, “Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm,” *Int. J. Comput. Sci. Mob. Comput.*, vol. 5, no. 2, pp. 24–33, 2016.
- [14] P. Yildirim, “Chronic Kidney Disease Prediction on Imbalanced Data by Multilayer Perceptron: Chronic Kidney Disease Prediction,” in *Proceedings - International Computer Software and Applications Conference*, 2017, vol. 2, pp. 193–198, doi: 10.1109/COMPSAC.2017.84.
- [15] E. H. A. Rady and A. S. Anwar, “Prediction of kidney disease stages using data mining algorithms,” *Informatics Med. Unlocked*, vol. 15, pp. 1–7, Jan. 2019, doi: 10.1016/j.imu.2019.100178.