International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

A Machine Learning Approach for the Diagnosis of Chronic Kidney Disease

Divya Pogaku[1], Sneha Bohra[2]

[1]PG Scholar, Department of Computer Science & Engineering, G H Raisoni University, Amravati, Maharashtra, India
Email ID: *pdivya565[at]gmail.com*

[2]Professor, Department of Computer Science & Engineering, G H Raisoni University, Amravati, Maharashtra, India

Abstract: ***Several billion people worldwide are afflicted by chronic kidney disease (CKD), a common and possibly fatal ailment. For efficient management and treatment of CKD, an early and precise diagnosis is essential. Machine learning (ML) algorithms have recently demonstrated promising results in a number of medical fields. This work uses a big dataset of clinical, demographic, and laboratory data gathered from a large cohort of CKD patients to provide a machine learning method to CKD diagnosis. To ensure the dependability and robustness of the ML models, the dataset handle values that are missing, outliers, and class imbalances during pre - processing. Numerous machine learning techniques encompass k - nearest neighbors (KNN), Random Forest, logistic regression, support vector machines (SVM), Naive Bayes classifiers, and feed - forward neural networks. Cross - validation technique is applied to train and test feed forward neural networks. Model performance is evaluated by estimating the accuracy of the model, with the Random Forest machine learning model achieving the highest accuracy. With the use of perceptron, we suggested a combined model that combines random forest with logistic regression which has the best accuracy; as a result, we hypothesized that more complicated clinical data may be used with this technology to diagnose disorders.***

Keywords: KNN, Naive Bayes classifier, Logistic Regression, Random Forest and SVM

1. Introduction

Chronic kidney disease (CKD) is a prevalent medical condition that involves the progressive and permanent decline in kidney function. It poses a significant global health challenge, impacting a large population. If left unmanaged, CKD can give rise to a range of complications and greatly diminish an individual's quality of life. Chronic kidney failure, a global public health concern, impacts approximately 10% of the world's population. In China, the prevalence of CKD is approximately 10.8%, while in the United States, it ranges from 10% to 15%. A survey conducted on Mexican population stipulate that 14.7% of population suffering from this kind of issue. Chronic kidney failure is a global public health concern which is affecting 10% the world's population. Chronic kidney failure is a pathologic weakness which cannot be cured and this could also lead to cardiovascular diseases.

In the field of medicine, machine learning has been employed to identify different ailments, examine critical disease - related factors, and detect anomalies in the human body. Regressiontree (RT), probability, decision surface, and a classification based on a neural network were used in these models. Polat et al. developed a super vector machinedepends on the method used to choose the features. The accuracy of the proposed models ranged from 97.75% to 98.5%. Aljaaf, J., et al. filled in the missing data using innovative multiple imputation. MLP has the accuracy of 99.75% in the model was developed by Boukenze et al. Studies usually concentrate on creating models and producing desired outcomes. There isn't a detailed explanation of the entire procedure for filling in the missing numbers, and there isn't any technology to help choose predictors based on features. Almansour et al. has accomplish an accuracy of 97.75% using various models. Gunarathne et al., has used zero to fill missing data and 99.1% accuracy is achieved by using decision forest. Moreover, the majority of the time was spent using earlier studies, which was available through the UCI's machine learning repository. Due to this, their strategy is insufficient whenever test results for samples are unclear. In actuality, patients may for a variety of reasons forego several tests before arriving at a diagnosis.

2. Literature Review

FuRESs and FOAM, two internal fuzzy classifiers, were examined to see if they might be used to diagnosing individuals with chronic kidney disease (CKD). A linear classifier, was utilized for comparison. The Machine Learning Repository provided the CKD data utilized in this study. To test the resilience of the two fuzzy techniques, composite dataset was made by addition of various amounts of proportional noise. After adding 11 stages of proportional noise sequentially on every numerical the training and prediction sets, and the attribute were initially mixed in pairs. In order to compare the categorization rates for these 121 couples, a grid was created containing 121 sets of simulated data. Second, using simulated the two fuzzy classifiers' performances were evaluated on dataset with 11 stages of noise distributed at random to every numeric attribute. FuRES and FOAM have 200 bootstrap Latin partitions. The average prediction rates of 98.1 0.5% and 97.2 1.2%, respectively. With the same evaluation, the PLS - DA may provide 94.3 0.8%. FuRES, FOAM, and PLS - DA classifiers models were also assessed using confluent datasets made up to the original dataset and modified dataset. The 200 - bootstrapped average predicted rate for FuRES and FOAM findings show that both FuRES and FOAM are effective in identifying CKD patients. These two fuzzy categories are crucial tools for the accurate identification of individuals with CKD. [1].

We conducted surveying a cross - section of a sample for Chinese people individuals that was nationally representative to determine if albuminuria was present. Participants had their blood pressure checked, samples of blood and urine were obtained, and they completed a questionnaire on their lifestyle and medical history. GFR was calculated using measurements of serum creatinine. Compared to other locations, albuminuria was independently connected with economic growth in rural areas. Age, sex, diabetes, high blood pressure, hyperuricemia, location where the patient lives and also patient's financial status were additional traits that are linked to kidney impairment [2].

It is possible that the management of chronic illnesses may be significantly improved by predictive models created utilizing temporal data from electronic health records (EHRs). These data, however, come with a host of technical difficulties, such as erratic data collection and variable patient history duration. In this article, we outline and compare three alternative methods for applying machine learning to create prediction model using the patient's temporal EHR data. The first technique, this integrates the scores Together with the patient's medical history, one or more predictors and is a common non - temporal strategy. Anothertwo approaches take advantage of the temporal changes the data. The two temporal methods handle missing data differently and represent temporal information differently. We created and evaluated a model to forecast changes in the most used kidney function indicator, the estimated glomerular filtration rate (eGFR), using information from the electronic health record (EHR) at Mount Sinai Medical Centre. Our findings suggest that temporal data might help predict kidney function decline by being included in the medical record of a person. They also show how crucial it is to include this knowledge in a certain way. Our findings, in particular, show that employed multi - task learning for the suitable technique to EHR data's temporal dynamics are captured robustly. Different predictors' relative values change with time [3].

The mortality rate can only be decreased by early discovery and efficient treatment. Machine learning algorithms are playing an increasing role in medical diagnostics because of them ability to categories data with high rates of accuracy. In order for classification algorithms to function correctly, appropriate feature selection techniques must be utilized to make datasets dimensions smaller. The category of Support Vector Machine technique was employing in this work to identify chronic renal disease. In order to shrink the chronic renal disease data file and identify the trouble, 2 important sorts of features strategies—methods using wrappers and filters—were applied. In a wrapper approach, both the classifier subset evaluator and the wrappers subset evaluate with the most effective the initial search engine were used. Stepwise The Greedy Search Engine and the Correlation - Based Evaluation of Features Subset Evaluator were used in the filter approach. consequently, the most effective the SVM classifier, First Search feature selection and the evaluator for filtered subsets methods achieved greater accuracy (98.5%), for the predicting a chronic kidney condition [4]. One of the most prevalent co - occurring conditions in people experiencing chronic kidney failure is anemia which is the end - stage renal disease. Agents that stimulate erythropoiesis (ESA), in particular, have emerged as the preferred therapy for that anemia. To uncover an appropriate medication for a patient, even the identical patients with varying Symptoms of anemia is quite difficult. This study builds upon previous research addressing the same issue, utilizing various techniques, including machine learning. However, this study specifically focuses on the population afflicted with chronic kidney disease (CKD) to examine their response to ESA/Iron therapy and develop a precise methodology for forecasting CKD. Furthermore, the three study - related nations—Spain, Italy, and Portugal—are represented in the ML method by both human and data inputs. A measure of the hemoglobin (Hb) prediction's mean absolute errors (MAE) for this model were close to or less than 0.6 g/dl, outperforming prior techniques [5].

Patients already suffering from chronic diabetes, major vascular disease and detected with early stage of chronic renal disease are at a risk to experience cardiovascular diseases. It is unknown how early chronic kidney disease affects borderline diabetes early - onset type 2 diabetic mellitus macrovascular outcomes. The inclusion of insulin had no difference in cardiovascular outcomes from usual treatment in the ORIGIN research (Outcome Reducing with an Initially Glargine Intervention). In this ORIGIN post hoc study, we assessed the cardiovascular results. Across participants with and without mild or methods for mild chronic renal illness. The evaluation of two co - primary composites cardiovascular outcomes. The first was the composite outcome of nonfatal MI death from cardiovascular cause; andthe second occurred when any of these occurrences were added to a revascularization surgery or a hospital stay for heart failure. Microvascular results, incidence diabetes, hypoglycemia, weight, and malignancies were among the pre - specified secondary outcomes. Patients with mild to severe renal disease experience a higher occurrence of combined primary ailments, such as cardiovascular mortality, nonfatal myocardial infarction, or nonfatal stroke, compared to patients without CKD [6].

Machine learning approach is used to detect Lung cancer by S Mukherjee, SU Bohra - The examination of the lungs has captivated medical experts throughout history, and even in modern times, it remains an intriguing field of investigation. In order to address this challenge, the development of a predictive system holds promise in reducing the threat to human life by enabling early detection of malignant growths. Numerous frameworks have been proposed, with many still in the experimental phase. One approach involves utilizing image data to detect cancerous cells, employing a neural network model to improve performance. A lung cancer prediction framework has been developed based on AI and deep neural networks, relying on supervised learning to achieve enhanced accuracy [7].

The diagnosis of medical conditions poses a significant concern for healthcare professionals, as critical decisions and treatment plans depend on accurate assessments. To address this challenge, a proposed framework aims to detect lung malignancy at an early stage through a two - phase process. The framework involves multiple steps, including image extraction, pre - processing, binarization, thresholding, division, feature extraction, and neural network identification. This model unfolds a Lung Cancer detection system based on machine learning and neural networks. Early detection of this malignancy plays a crucial role in saving patients' lives. In recent times, there has been significant growth in data mining and machine learning techniques for predicting various chronic diseases. In this particular model, CT scan images are utilized as inputs to predict the probability of the disease and its stages. A quick intelligent clinical decisions are framed to help medical practitioners to find out the disease in primitive stages and thereby makes the treatment cheaper. This machine learning study is used in this model to detect lung cancer and its stage detection which is propounded by S Mukherjee, S Bohra in 2020 [8].

Data mining techniques are efficient in detection of disease supporting medical predictions while reducing the burden on medical practitioners. In this model the primitive discovery of diabetes is done by using data mining. There are number of diseases linked to diabetes mellitus which are related to kidney, eye, and heart and with forth highest fatality rate in the world. The current study focuses on such issues by using data mining techniques accurate medical prediction is achieved. Chronic disease diabetes mellitus affects different body parts differently. Beforehand prediction of the disease will helps to save lives. Historically, the diagnosis of diabetes has relied on a battery - operated physical examination, which, in fact, lacks accuracy. To address this limitation, the current model utilizes Data Mining techniques, including Naive Bayes, K - Nearest Neighbor, Support Vector Machine, and Decision Tree, to predict the occurrence of diabetes mellitus. This detection model is developed by PG Palanimani, V Suresh Kumar, Sanket B Kasturiwala, ShafaqueAhmareen, SnehaBohra. [9].

3. Problem Definition

The renal function gradually deteriorates until it is completely lost. Symptoms of CKD do not appear till later. The ailment was therefore not identified until the kidney had around 25% of its original function. Timely identification, making lifestyle adjustments, and implementing suitable medical interventions are crucial for impeding the advancement of the disease, alleviating symptoms, and averting complications linked to CKD. The management of CKD centers on the objectives of impeding the progression of the disease, managing symptoms, and averting complications. Treatment plans are tailored to each individual and may encompass lifestyle adjustments, dietary modifications, medication administration, and addressing underlying conditions such as diabetes or high blood pressure. In advanced stages of CKD, life - sustaining interventions such as dialysis or kidney transplantation may be required which is expensive. To prevent this unfavorable state, early identification of CKD is necessary, as it enables healthcare professionals and patients to promptly implement appropriate medical interventions. Since this disease lacks evident symptoms, such predictive capability of this kind of Machine learning models holds the potential to save numerous lives by enabling early detection.

4. Methodology

In this part, multiple Machine Learning Methods are proposed. In order to analyze a data samples, multiple MLinnovation first constructed the classifiers. As potential components, the models that performed the best were picked. Investigating their miscalculations in judging the identification of the constituent models. Then, a desegregate model was constructed in order to get greater presentation.

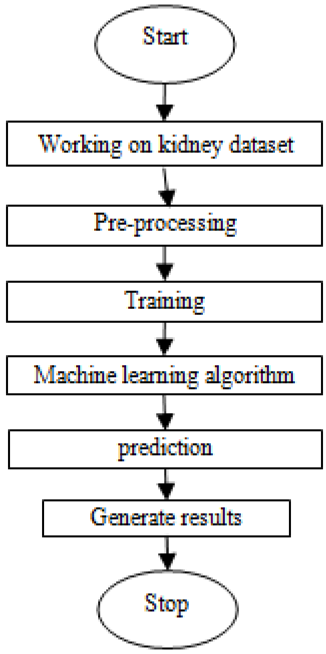


Figure 1: Proposed Method Block Diagram

1) Logistic Regression:

Logistic regression, a widely utilized algorithm in Machine Learning, falls under the category of Supervised Learning. This statistical method for solving binary classification problems aims to predict the probability of an event based on input factors. Because it models the link between the input variables and the event probability using the logistic function, it is known as a "logistic" model. In logistic regression, the input variables are linearly mixed with weights, and the logistic function is then applied to the outcome. Any real - valued number may be converted by the logistic function to a value between zeros and ones, which represents the likelihood that the event will occur. After then, a binary choice is made using this probability, often by using a threshold number. The logistic regression model is trained by maximizing the probability of the observed data, which entails obtaining what the weights' ideal values should be. Usually, optimization procedures like gradient descent are used for this. A few benefits of logistic regression are its simplicity, interpretability, and effectiveness. It is component of operate category and uninterrupted input variables together with can be developed to address multi - class classification challenges. It may not always be true in practice for the belief that there's a straight line among the variables being used with the event's log - odds to stay true.

2) Support Vector Classifier:

A well - liked machine learning technique The Support Vector Classifier (SVC) may be used for classification. It belongs to the family of support vector machine (SVM) algorithms. The objective of SVC is to find ideal hyperplane in a high - dimensional space which divides several classes of data points. SVC's main principle is to raise the size of the input feature space, which makes it easier to find a hyperplane that can split the classes. The approach determines the optimized increase in margin, which separates the hyperplane and the closest information points within every category. By utilizing a method known as the kernel trick, SVC is able to handle both linearly separable and non - linearly separable data. The kernel hack enables way to implicitly work in a higher - dimensional space without explicitly computing the changed characteristics. The SVC method works well with small to medium - sized datasets and does well when there are more characteristics than samples. Applications like text categorization, picture classification, and bioinformatics all benefit greatly from its utilization. SVC may have trouble with datasets that contain overlapping classes or noisy data, and its training time can be lengthy for large datasets. Additionally, selecting the proper kernel and fine - tuning hyper parameters might be difficult. However, SVC may be a potent tool for classification jobs when the right parameters are chosen and the data are handled carefully.

3) Random Forest:

Random Forest, a popular Machine Learning model, leverages the power of ensemble learning by incorporating decision trees. It excels in both classification and regression tasks. In a Random Forest, multiple decision trees are built using random subsets of data, with each tree trained on a distinct portion of the dataset, employing feature selection. Based on the combined results of all the different decision trees, the algorithm makes predictions. A majority vote (for classification) or average (for regression) of the predictions from all the trees determines the final classification or prediction for the target variable. Random Forests are renowned for being reliable, scalable, and capable of handling high - dimensional data with intricate interactions. Additionally, they have internal defenses for coping with handling outliers and missing values. The Random Forests technique is more resistant to noise and volatility in the data and helps to prevent over fitting. Due to the algorithm's ability to calculate each feature's contribution to total predictive power, it also enables feature significance analysis. All things considered, Random Forests are a flexible and strong algorithm that are frequently applied in many industries, including banking, healthcare, and natural language processing.

4) K - Nearest Neighbors:

An algorithm for classification and regression problems includes KNN. It is a lazy training model, which means that it does not assume anything about the distribution of the data's underlying values and does not demand any training before producing predictions. In order to forecast the class label of a new input data point, the KNN method first determines the K nearest data points to it. It then uses the class labels of those closest neighbors. K is a hyper parameter that may be altered dependent on the current problem. If K is greater, the choice boundary will be smoother; if K is less, the decision border will be more complicated and limits decision - making. The KNN method predicts the output value for a new data point as the average of that point's K closest neighbours' output values in regression tasks. KNN is a straightforward and understandable algorithm that can be effective for some kinds of issues. However, for big datasets, it can be computationally costly and necessitates careful feature scale and selection. Furthermore, the algorithm may be vulnerable to irrelevant and noisy information, which might result in subpar performance.

5) Naïve Bayes:

For classification tasks, Naive Bayes, an intelligence learning algorithm, is used. The Bayes theorem is used to support the premise that characteristics if the class and conditional independence. The algorithm assesses the probability of each class given the input attributes and chooses the most likely class as the forecast. In problems involving text categorization and other areas where the independence assumption is true, the Naive Bayes method performs well. It is a probabilistic method that works with categorical and continuous data. In addition to being computationally effective, Naive Bayes is capable of handling big datasets with high dimensionality. The method operates by first deriving from the likelihood of every characteristicgiven each class in the training data. Then, given the input characteristics, the chance of each class is calculated using the Bayes theorem, and the group with the greatest probability is selected as the forecast.

Applications including spam filter, sentiment analyzing, and document categorization have all shown success using naive Bayes. It is a potent tool for creating straightforward yet efficient classifiers, and both business and academics make extensive use of it.

5. Results and Discussions

Home Page: Viewers may access the front page of the website A MLmethod for Diagnosiechronic kidney failure here.

Figure 2: Home page

About: On the about page, users is fecilitaed to go throughinformation about Machine Learning (ML) technique to Diagnose Chronic RenalFailure and also the symptoms of the issue.

Paper ID: SR23609152846 DOI: 10.21275/SR23609152846

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

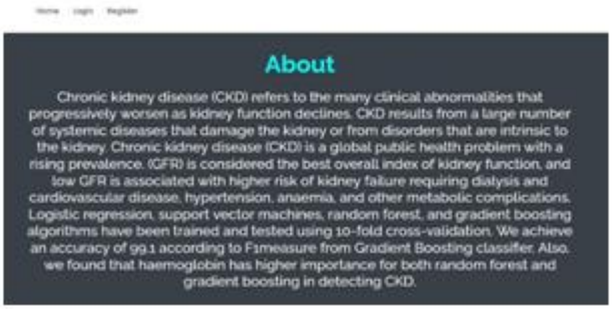


Figure 3: About page

Register: Users can register for this programmed, which uses machine learning to diagnose chronic kidney disease (CKD).

Figure 4: Register page

Login: A ML (Machine Learning) authentication page for a chronic kidney disease (CKD) diagnosis.

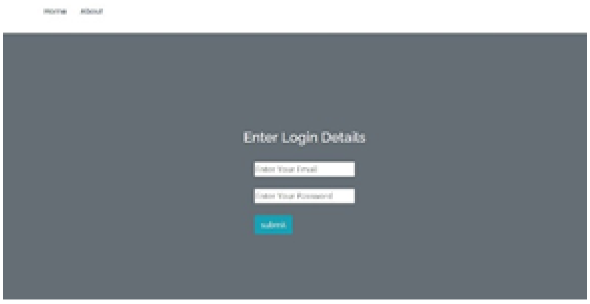


Figure 5: Login

Login Homepage: A ML Methodology for Diagnosing Chronic Kidney (CKD) Disease's primary login page for users.

Figure 6: Login Homepage

Upload data set: The user uploads his dataset in this section of the upload page.

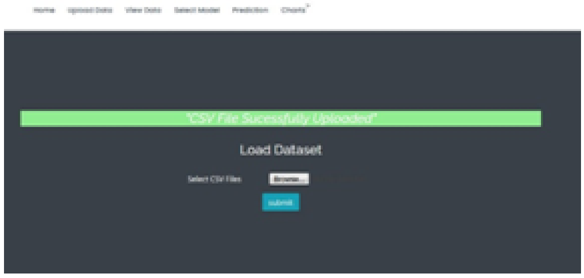


Figure 7: Upload data set

View data set: The user is able to view the dataset that they submitted.

S/N	id	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	bu	sc	sod	pot	hemo	pcv	wc	rc	htn	cad	a
1	0	48.0	80.0	1.02	1.0	0.0	nan	normal	notpresent	notpresent	121.0	36.0	1.2	nan	nan	15.4	44.0	7800.0	5.2	yes	no	g
2	1	7.0	50.0	1.02	4.0	0.0	nan	normal	notpresent	notpresent	nan	18.0	0.8	nan	nan	11.3	38.0	6000.0	nan	no	no	g
3	2	62.0	80.0	1.01	2.0	3.0	normal	normal	notpresent	notpresent	423.0	53.0	1.8	nan	nan	9.6	31.0	7500.0	nan	no	no	p
4	3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	56.0	3.8	111.0	2.5	11.2	32.0	6700.0	3.9	yes	no	p
5	4	51.0	80.0	1.01	2.0	0.0	normal	normal	notpresent	notpresent	106.0	26.0	1.4	nan	nan	11.6	35.0	7300.0	4.6	no	no	g
6	5	60.0	90.0	1.015	3.0	0.0	nan	nan	notpresent	notpresent	74.0	25.0	1.1	142.0	3.2	12.2	39.0	7800.0	4.4	yes	no	g
7	6	68.0	70.0	1.01	0.0	0.0	nan	normal	notpresent	notpresent	100.0	54.0	24.0	104.0	4.0	12.4	36.0	nan	nan	no	no	g
8	7	24.0	nan	1.015	2.0	4.0	normal	abnormal	notpresent	notpresent	410.0	31.0	1.1	nan	nan	12.4	44.0	6900.0	5.0	no	no	g
9	8	52.0	100.0	1.015	3.0	0.0	normal	abnormal	present	notpresent	138.0	60.0	1.9	nan	nan	10.8	33.0	9600.0	4.0	yes	no	g
10	9	53.0	90.0	1.02	2.0	0.0	abnormal	abnormal	present	notpresent	70.0	107.0	7.2	114.0	3.7	9.5	29.0	12100.0	3.7	yes	no	p
11	10	50.0	60.0	1.01	2.0	4.0	nan	abnormal	present	notpresent	490.0	55.0	4.0	nan	nan	9.4	28.0	nan	nan	yes	no	g
12	11	63.0	70.0	1.01	3.0	0.0	abnormal	abnormal	present	notpresent	380.0	60.0	2.7	131.0	4.2	10.8	32.0	4500.0	3.8	yes	no	p

Figure 8: View data set

Training: We're training the algorithm to see which one has the best accuracy.

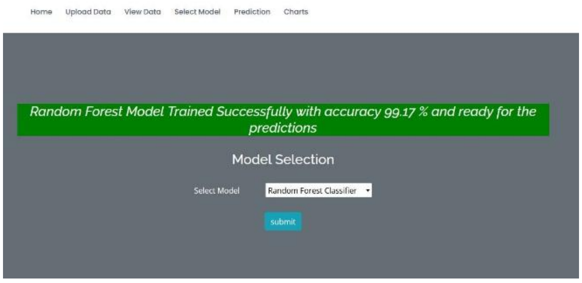


Figure 9: Random Forest Model Selection

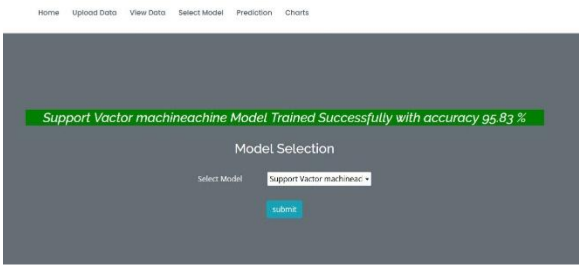


Figure 10: SVM Model Selection

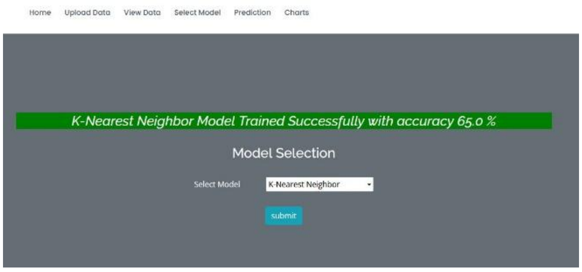


Figure 11: KNN Model Selection

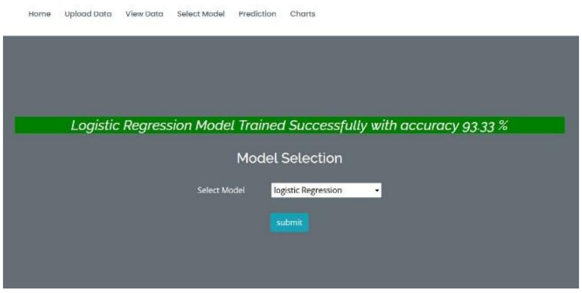


Figure 12: Logistic Regression Model Selection

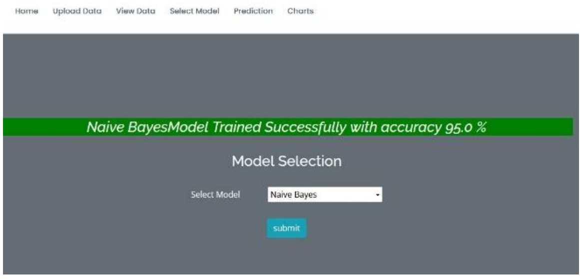


Figure 13: Naive Bayes Model Selection

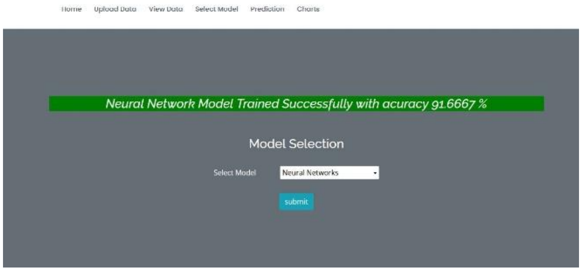
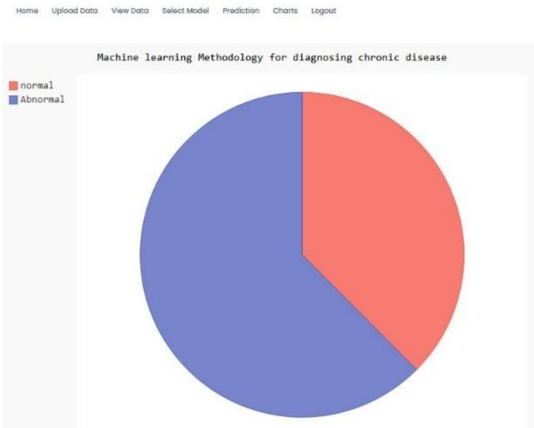


Figure 14: Neural Network Model Selection

Detection: The patient's results for the identification of chronic renal disease are shown on this page.

Figure 15: Treatment is necessary for the patient's chronic renal illness

Charts: The patient's record is detection results for chronic renal disease is displayed on this page.



Test cases:

S. No	Test cases	I/P	Expected O/P	Actual O/P	P/F
1	Read datasets.	Datasetpath.	Datasets to read.	Datasets fetched successfully.	It produced P. If this not F will come
2	Verifying the chronic kidney disease identify type of disease	Input for chronic kidney disease classification	The output can be classified as either chronic kidney disease (CKD) or non - chronic kidney disease (CKD).	Output is classified as chronic kidney disease	It produced P. If this is not, it will undergo F
3	Verifying the chronic kidney disease identify type of disease	Input for chronic kidney disease classification	The output is categorized as either chronic kidney disease (CKD) or the absence of chronic kidney disease (non - CKD).	Output is classified as Not chronic kidney disease	It produced P. If this is not, it will undergo F
4	Verifying the chronic kidney disease identify type of disease	Input chronic kidney disease for prediction the disease	Need to predict the best accuracy	Model successfully predicted best accuracy	It produced P. If this is not, it will undergo F

6. Conclusion

The recommended CKD diagnostic technique is feasible with regard to sample diagnosis and data imputation. After KNN replacement was used to automatically fill the values that were missing in the data set, the integrated model was able to achieve a suitable degree of accuracy. We thus hypothesis that using this approach to really diagnose CKD would have a positive outcome. Furthermore, incorporating clinical information from other diseases that are typically used in real - world medical diagnoses could potentially enhance the effectiveness of this technique. However, it is important to note that the model's generalization performance might be limited due to the scarcity of available data samples during the model - building process, caused by restrictions in the circumstances. Additionally, it should be emphasized that the model is not designed to evaluate the severity of CKD based on the data samples from the two groups (CKD and non - CKD).

References

[1] Z. Chen et al., “Diagnosis of patients with chronic kidney disease by using two fuzzy classifiers, ” Chemometr. Intell. Lab., vol.153, pp.140 - 145, Apr.2016.

[2] L. Zhang et al., “Prevalence of chronic kidney disease in china: a cross - sectional survey, ” Lancet, vol.379, pp.815 - 822, Aug.2012.
[3] Singh et al., “Incorporating temporal EHR data in predictive models for risk stratification of renal function deterioration, ” J. Biomed. Inform. vol.53, pp.220 - 228, Feb.2015.
[4] H. Polat, H. D. Mehr, A. Cetin, “Diagnosis of chronic kidney disease based on support vector machine by feature selection methods, ” J. Med. Syst., vol.41, no.4, Apr.2017.
[5] C. Barbieri et al., “A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease patients undergoing dialysis, ” Comput. Biol. Med., vol.61, pp.56 - 61, Jun.2015.
[6] V. Papademetriou et al., “Chronic kidney disease, basal insulin glargine, and health outcomes in people with dysglycemia: The origin study, ” Am. J. Med., vol.130, no.12, Dec.2017.
[7] S. Mukherjee and S. U. Bohra, "Lung Cancer Disease Diagnosis Using Machine Learning Approach, " 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp.207 - 211, doi: 10.1109/ICISS49785.2020.9315909.
[8] S Mukherjee, S Bohra “MACHINE LEARNING APPROACH FOR LUNG CANCER DETECTION AND STAGES PREDICTION”. Journal of Emerging Technologies and Innovative Research, vol.7, issue 1, Jan.2020.
[9] PG Palanimani, V Suresh Kumar, Sanket B Kasturiwala, ShafaqueAhmareen, SnehaBohra, A novel approach for detection of diabetes patients using machine learning techniques in International Interdisciplinary Humanitarian Conference for Sustainability (IHC), 1402 - 1408, 2022.
[10] S. Parveen, M. Shahbaz, A. Guergachi, and K. Keshavjee, Performance Analysis of Data Mining Classification Techniques to Predict Diabetes, ProcediaComput. Sci., vol.82, no. Supplement C, pp.115 - 121, Jan.2016.
[11] U. N. Dulhare and M. Ayesha, Extraction of action rules for chronic kidney disease using Naive Bayes classifier, in 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2016, pp.15.
[12] N. Borisagar, D. Barad, and P. Raval, Chronic Kidney Disease Prediction Using Back Propagation Neural Network Algorithm, in Proceedings of International Conference on Communication and Networks, Springer, Singapore, 2017, pp.295 - 303.
[13] Pritom, M. A. R. Munshi, S. A. Sabab, and S. Shihab, predicting breast cancer recurrence using effective classification and feature selection technique, in 2016 19th International Conference on Computer and Information Technology (ICCIT), 2016, pp.310 - 314.
[14] S. Mishra, P. Chaudhury, B. K. Mishra, and H. K. Tripathy, An Implementation of Feature Ranking Using Machine Learning Techniques for Diabetes Disease Prediction, in Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, New York, NY, USA, 2016, p.421 - 423.
[15] D. Zufferey, T. Hofer, J. Hennebert, M. Schumacher, R. Ingold, and S. Bromuri, Performance comparison of multi - label learning algorithms on clinical data for chronic diseases, Comput. Biol. Med., vol.65, no. Supplement C, pp.34 - 43, Oct.2015.
[16] S. Bashir, U. Qamar, F. H. Khan, and M. Y. Javed, MV5: A Clinical Decision Support Framework for Heart Disease Prediction Using Majority Vote Based Classifier Ensemble, Arab. J. Sci. Eng., vol.39, no.11, pp.7771 - 7783, Nov.2014.
[17] Lee, Hyung - Chul, et al. "Derivation and validation of machine learning approaches to predict acute kidney injury after cardiacsurgery. " Journal of clinical medicine 7.10 (2018): 322.
[18] Mohamadlou, Hamid, et al. "Prediction of acute kidney injury with a machine learning algorithm using electronic health record data. " Canadian journal of kidney health and disease 5 (2018): 2054358118776326.