Abstract :

Early kidney disease prediction using machine learning is a method of identifying individuals at high risk for developing kidney disease by training a model on a dataset's of patient information and lab test results. The goal is to detect the disease in its early stages before significant damage has occurred, allowing for early intervention and treatment. This can be done through various machine-learning classification algorithms. The model is trained on a lab test results, and then used to predict the risk of kidney disease in new patients. This approach has the potential to improve patient outcomes by detecting the disease earlier and reducing healthcare costs. Additionally, machine learning can also be used to identify patterns and correlations in patient data that might not be immediately apparent to healthcare providers. This can help in understanding the disease better and in identifying the risk factors for kidney disease. Furthermore, it can also help in identifying the patient who needs more attention and care, which can lead to better patient outcomes.

Keywords : Machine learning , Artificial neural network , Chronic kidney disease ( CKD ),

Logistic regression (LR) , Random forest (RF) , K-nearest neighbors (KNN) .

Introduction :

Related work :

In recent years, studies on the diagnosis of chronic kidney disease has been conducted. There are many techniques to forecast CKD. An artificial neural network (ANN)-based software tool was designed to identify the status of patients by T. Di Noia et al. [1] The software tool is available as a web application and an Android application. Four machine learning techniques were used, and the classifier was chosen that gave the best results after being analysed with a dataset containing details from Indians at Apollo Hospital. [2] The accuracy achieved was 98.3%, and the best classifier was SVM.

As per Electronic Health Records (EHR), patients are classified into two types:

117 progressor patients and 364 non-progressor patients based on

glomerular filtration rate (GFR) GFR )[3].A predictive model was developed using

Nave Bayes and logistic regression classifiers Metabolic features are compared in both types. The results reveal that the phosphate values are higher in progressors compared to non-progressors.

Sahil Sharma, Vinod Sharma, and Atul Sharma used a dataset containing 24 attributes and classified it using different techniques. To determine the accuracy, they compared the actual results to the predicted results. Performance evaluation criteria include accuracy, sensitivity, and precision. The results revealed that the technique had a nearly perfect accuracy of 98.6%, a sensitivity of 0.9720, and a precision of 1 [4]. K. A. Padmanaban and G. Parthiban performed a study on identifying CKD in diabetic patients. 600 records are taken, and machine learning methods like decision trees and Naive Bayes are used. The decision tree algorithm achieved a higher accuracy of about 91% [5].

Muhammad Minoar Hossain analysed several machine learning algorithms and used a max-voting ensemble model with five classifiers. A dataset having 400 medical records and 24 features is used. Two types of data are used: nominal and numerical. Linear discriminant analysis is the feature optimization technique used. The accuracy of feature selection was approximately 99.5%. Six feature optimization techniques are analysed to identify the strengths of feature optimization. [6]

References :

1.Almasoud, M., & Ward, T. E. (2019). Detection of chronic kidney disease using machine

learning algorithms with least number of predictors. *International Journal of Soft*

*Computing and Its Applications*, *10*(8).

2.Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of

machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159.

3. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., & Cho, H. (2015). Xgboost:

extreme gradient boosting. *R Package Version 0.4-2*, *1*(4), 1–4.

4. Chittora, P., Chaurasia, S., Chakrabarti, P., Kumawat, G., Chakrabarti, T., Leonowicz, Z.

. . ., et al. (2021). Prediction of chronic kidney disease-a machine learning

perspective. *IEEE Access*, *9*, 17312–17334.

5. Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network

classification models: a methodology review. *Journal of Biomedical Informatics*,

*35*(5–6), 352–359.

6. Ghosh, P., Shamrat, F. J. M., Shultana, S., Afrin, S., Anjum, A. A., & Khan, A. A.

(2020). Optimization of prediction method of chronic kidney disease using machine

learning algorithm. In *2020 15th International joint symposium on artificial intelligence*

*and natural language processing* (pp. 1–6). IEEE.

7. Gunarathne, W. H. S. D., Perera, K. D. M., & Kahandawaarachchi, K. A. D. C. P. (2017).

Performance evaluation on machine learning classification techniques for disease

classification and forecasting through data analytics for chronic kidney disease

(CKD). In *2017 IEEE 17th International conference on bioinformatics and bioengineering*

(pp. 291–296). IEEE.

8. Tharwat, A., Gaber, T., Ibrahim, A., & Hassanien, A. E. (2017). Linear discriminant

analysis: A detailed tutorial. *AI Communications*, *30*(2), 169–190.

9.

1. T. Di Noia *et al*, "An end stage kidney disease predictor based on an artificial neural networks ensemble," *Expert Syst. Appl.,* vol. 40, *(11),* pp. 4438-4445, 2013.
2. Anusorn Charleonnan, Thipwan Fufaung, Tippawan Niyomwong,

Wandee Chokchueypattanakit, Sathit Suwannawach, Nitat

Ninchawee, “Predictive Analytics for Chronic Kidney Disease Using

Machine Learning Techniques,” Proc. Management and Innovation

Technology International Conference (MITiCON-2016) , IEEE, Oct. 2016

1. H. S. Chase *et al*, "Presence of early CKD-related metabolic complications predict progression of stage 3 CKD: a case-controlled study," *BMC Nephrology,* vol. 15, *(1),* pp. 187, 2014..
2. Sahil Sharma, Vinod Sharma, Atul Sharma, “Performance Based

Evaluation of Various Machine Learning Classification Techniques

for Chronic Kidney Disease Diagnosis,” July18, 2016.

1. K. A. Padmanaban and G. Parthiban, "Applying Machine Learning Techniques for Predicting the Risk of Chronic Kidney Disease," *Indian Journal of Science and Technology,* vol. 9, *(29),* 2016.
2. “Analysis of the performance of feature optimization techniques for the

diagnosis of machine learning-based chronic kidney disease , 2022”,

ScienceDirect.