AUTOMATED CHRONIC KIDNEY DISEASE DETECTION SYSTEM

Ratna[1], Shreya Bapu Gadage[2], Smruti Timmanna Nayak[3], Uma G S[4], Dr. Kanakaraju R[5]
[1,2,3,4,5]Department of CSE, Don Bosco Institute of Technology, Bengaluru, Karnataka, India
patilprutvi0@gmail.com, gadageshreya134@gmail.com, smrutinayak281@gmail.com, umagsumags2004@gmail.com, kanakaraju.r@dbit.co.in

Abstract

Alcohol abuse, industrial pollution, global warming, drug abuse, and the effects of dangerous lifestyles are some of the causes contributing to the growth in kidney failure. These factors all play a part in the rising incidence of renal diseases and the strain on healthcare systems. If early detection by machine learning (ML) is feasible, the clinical load is decreased. This study examines many machine learning techniques for the categorization and prediction of renal disorders. To choose the models that achieve the maximum accuracy for the specified dataset, a thorough comparison of over six models is conducted. Although ensemble approaches and hyperparameter tuning boost accuracy somewhat, they come at the expense of more calculation time, negating their practical impact. This study focuses on practical ML-related approaches toward offline CKD detection with primary concern on the best accuracy for diagnostic application.

Keywords: *Kidney Failure, Renal Diseases, Health Systems, Early Detection, Machine Learning, Prediction, Accuracy, CKD Detection, Diagnostic Application*

Introduction

The chronic kidney disease (CKD) is a widespread, progressive disorder affecting millions worldwide, characterized by the gradual loss of kidney function over time. One of the main causes of illness and death in the globe is chronic kidney disease (CKD). According to recent estimates, 10% of people globally will experience renal sickness at some point in their lives. Since the condition typically shows no symptoms in its early stages, early detection and treatment are difficult. Prior to diagnosis, patients may lose up to 25% of their kidney function, suggesting that serious and often irreversible damage may have already occurred by the time symptoms manifest. This delayed identification is the primary cause of the disease's progression to end-stage renal disease (ESRD), where the only viable treatments are dialysis or a transplant. A combination of environmental, genetic, and lifestyle factors can lead to chronic kidney disease.

The kidneys, which are composed of millions of filtering units called nephrons, eventually lose their ability to filter waste and preserve fluid and electrolyte balance as a result of chronic stress. In addition to kidney failure, chronic kidney disease (CKD) significantly increases the risk of cardiovascular diseases and other outcomes. Furthermore, the ignorance of at-risk persons might exacerbate the illness, especially in disadvantaged regions where access to professional diagnostic services and medical care is limited. It is one of the most challenging disorders to treat because of the lack of clear, early-stage symptoms, which delays diagnosis and results in poor prognoses.

The high expense of diagnosis, the scarcity of nephrologists in rural areas, and the dependence on invasive diagnostic tests make conventional healthcare approaches insufficient. In order to facilitate risk assessment, disease monitoring, and early identification, it is increasingly imperative to incorporate clever computational tools. In this sense, artificial intelligence (AI) and machine learning (ML) have emerged as promising strategies to transform CKD diagnoses by analyzing complex medical information to identify early sickness

indications.

According to recent research, machine learning algorithms including Support Vector Machines (SVM), Artificial Neural Networks (ANN), Decision Trees, and k-Nearest Neighbors (kNN) can predict CKD with high accuracy.. These models use clinical features like blood pressure, protein levels in urine, glucose levels, and body mass index to classify and predict disease status. In contrast to traditional methods, which may suffer from human error and are time-consuming, ML models can process large volumes of patient data efficiently, learning from patterns and improving predictive capabilities over time. Such techniques offer the dual benefit of reducing diagnostic cost and providing rapid assessments, especially beneficial in resource-constrained settings.

Literature Review:

Narinder Kumar and Sanjay Singla (2024):

Chronic kidney disease prediction: Optimization of machine learning algorithms- [1] This paper discusses an investigation on different machine learning techniques used to predict the early stage of kidney disease from the data of patients. The study is executing six algorithms and relatively choosing a more accurate one. It also points out the conflicts between precision and performance consumption and offers a solution for the appropriate models for non-internet and real healthcare environments.

Sonam Bhandurge, Dr. Kuldeep Sambrekar, Dr.Shridhar Allagi, Dr. Vijay S.Rajpurohit and Dr. Sanjeev Sannakki (2024):

A Proposed Model for Machine Learning-Based Early Chronic Kidney Disease Detection- [2] Paper offers a machine-based model as an apt way for the early unveiling of chronic kidney disease (CKD) with the use of supervised algorithms i.e. SVM, Random Forest, ANN, and Logistic Regression. The matter deals with preprocessing, imputation, and feature selection. One of the algorithms, Random Forest, gave a perfect prediction score of 100% which symbolizes that it is a powerful tool for the early diagnosis and treatment of CKD in clinical settings.

Garima Shukla, Gaurav Dhuriya, Sofia K Pillai and Aradhna Saini (2023):

Chronic Kidney Disease Prediction Using Machine Learning Algorithms and the Important Attributes for the Detection- [3] Paper is on the use of machine learning techniques for the prediction of chronic kidney disease (CKD) at the very early stages. The methods employed were K-Nearest Neighbors, Decision Tree, and Artificial Neural Networks. The researchers explained that they looked into previous medical files and different sorts of tests, and then finally, they zeroed in on some of the key attributes that affect CKD detection. The Decision Tree algorithm yielded an accuracy rate of 98.60%. The study addresses the role of supervised learning models and gives a thorough exploration of features and metrics that are most important to the work of CKD diagnosis and prevention.

Rama Al-Momani, Ghada Al-Mustafa, Razan Zeidan, Hiam Alquran, Wan Azani Mustafa and Ahmed Alkhayyat (2022):

Chronic Kidney Disease Detection Using Machine Learning Technique- [4] It investigates early detection of chronic kidney disease (CKD) through the application of machine-learning algorithms Artificial Neural Networks, Support Vector Machines, and k-Nearest Neighbors. Through using a dataset of 400 records with 13 features, the research suggests that ANN gives the best results in terms of accuracy (99.2%), commenting on the power of AI in a very clear and stringent manner where it could locate the disease much faster.

Pankaj Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawati, Tulika Chakrabarti, Michal Jasinski, Łukasz Jasinski and Radomir Gono (2021):

A Machine Learning Approach to Chronic Kidney Disease Prediction- [5] In the publication, the author describes a study that used the UCI CKD dataset to predict chronic kidney disease (CKD) using seven machine learning classifiers and one deep neural network. Several feature selection methods are used in the study, and the class imbalance issue was resolved by applying the Synthetic Minority Over-sampling Technique (SMOTE). The results showed that the performance of CKD prediction is significantly improved, with the Linear SVM with L2 penalty and full features giving the highest accuracy of 98.86% across all models, while the deep neural network reached as high as 99.6%.

M.S. Abirami, Sayak Das and Roopal Sood (2024):

Chronic Kidney Disease Identification using Random Forest and XGBoost- [6] Author Name mainly addresses the prompt detection of chronic kidney disease (CKD), which is a progressive illness that may lead to fatal events in the absence of a timely diagnosis. Nowadays, there is no standard multistage algorithm for prepared induction reasons in the early stages of the illness, and having this kind of guidance makes the task of choosing which subgroup of patients needs closer monitoring and early caring pretty difficult. Graph-based boosting (XGBoost) and Random Forest are the main methods for managing CKD cases that the research uses to classify the cases more accurately. The strategy is realised to give early intervention, thus, curing the patient of the disease and avoiding the use of advanced treatments that are as extreme as dialysis or transplantation. The findings of the study strongly point out that model performance is primarily subject to the quality and the reliability of the data set used.

Anurag, Narayan Vyas, Vishal Sharma and Deepak Balla (2023):

Chronic Kidney Disease Prediction Using Robust Approach in Machine Learning- [7] The article introduces a machine learning model as a tool to diagnose chronic kidney disease (CKD) using the data available from the patient's demographic, clinical, and laboratory records. The process combines a search for the most significant features, the reduction of input space, and decision-making through algorithms to find the most relevant risk parameters and thus a high-accuracy rate (89%). The case presents the capability of the advanced and smart diagnosis selections which are capable of adapting, are interactive, and are capable of not only learning from the newly input information but also of recommending and helping the healthcare provider in their decision-making of timely, accurate and informed treatment options for CKD.

B.Seetha Lakshmi, Vijay Anand R, Habeeb Omotunde, M. Birundadevi, Mahalakshmi S and Madala Guru Brahmam (2024):

Deep Learning for Automated Classification of Kidney Lesions in Chronic Kidney Disease Patients- [8] In this research, it presented a deep learning-based method for precise detection of renal abnormalities from quantitative as well as qualitative medical imaging information. The introduced model outperformed state-of-the-art methods in precision, recall, and diagnostic accuracy for chronic kidney disease (CKD). The model also overcame principal data privacy issues, providing a robust and ethical solution. This innovation has high potential for revolutionizing CKD diagnosis and treatment, and it calls for further studies in automated nephrology diagnostics.

Akash Maurya, Rahul Wable, Rasika Shinde , Sebin John , Rahul Jadhav and Dakshayani.R (2019):

Machine Prediction of Chronic Kidney Disease and Suggestion of an Appropriate Dietary Plan [9]. In this work, it introduces a deep learning system to classify renal abnormalities based on qualitative and quantitative

medical imaging data. The system, trained on a well-curated dataset, achieved better performance compared to recent approaches in high-accuracy and precision-based diagnosis of chronic kidney disease (CKD) and recall. It also meets ethical requirements for patient data privacy. This research provides a promising contribution to the field of medical image analysis and paves the way for better diagnostic devices in nephrology treatment.

Guozhen Chen, Chenguang Ding, Yang Li, Xiaojun Hu, Xiao Li, Li Ren, Xiaoming Ding, Puxun Tian, and Wujun Xue (2020):

Predicting Chronic Kidney Disease using an Adaptive Hybridized Deep Convolutional Neural Network on the Internet of Medical Things Platform-[10] The study's author recommends an Adaptive Hybridized Deep Convolutional Neural Network (AHDCNN) for the early detection and classification of chronic kidney disease, especially kidney cancer. The method improves classification accuracy and reduces feature size, allowing for effective tissue discrimination. Built on the Internet of Medical Things (IoMT) framework, the suggested system offers a dependable diagnostic tool through the use of deep learning and predictive analytics, opening the door for clever, AI-based nephrology healthcare solutions.

Dupati Lohith Chowdari, Kesamreddy Vinay Kumar Reddy, Shoba L K and Pongiannan R K (2025):

Optimizing Chronic Kidney Disease Detection through Multi-Modal Data Fusion using CNN-MLP Architectures- [11] In this paper, it introduces a Multi-Modal Learning system that incorporates clinical biomarkers with using medical imaging to improve chronic kidney disease detection by increasing accuracy. Using Convolutional Neural Networks, the system for image processing and Multi-Layer Perceptron (MLP) for structured data, implemented through a fusion layer. Advanced preprocessing provides better inputs, and the composite model minimizes diagnostic mistakes. This methodology provides a more consistent, computerized solution for early diagnosis of CKD, overcoming some of the limitations found in conventional techniques.

Vanathi.D, S.M. Ramesh, Sudha.K, Tamizharasu.K, Sengottaiyan.N and Kalyanasundaram.P (2024):

A Perspective on Using Machine Learning to Predict Chronic Kidney Disease- [12] By applying a variety of machine learning algorithms, the authors of this work hope to improve the prediction accuracy of chronic kidney disease (CKD). They suggest comparing Artificial Neural Networks (ANN), Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) in addition to four ensemble approaches: Random Forest, Extra Trees, AdaBoost, and XGBoost. Their findings demonstrated that the KNN, SVM, and ANN models performed better than the ensemble methods, with a high reliability of 99.2% for the early identification and diagnosis of chronic kidney disease.

Navaneeth Bhaskar and Suchetha M (2019):

A Model for Automated Chronic Kidney Disease Detection Based on Deep Learning- [13] By examining salivary urea, the author of this research proposes a novel sensing technique for the early detection of renal illness. A one-dimensional Convolutional Neural Network (CNN) in conjunction with a Support Vector Machine (SVM) classifier was used to decode raw sensor inputs. The classification accuracy was considerably improved to 98.04% by the CNN–SVM combination model. The proposed method highlights the promise of saliva-based biosensing for clinical diagnostics by providing an accurate, non-invasive, machine learning-based method for kidney disease diagnosis.

Espin Shalo S P, Jeeshan R J, S Gunasekar, Pius Agbulu G and Venkat Ramana N (2024):

Automated CKD Detection via Image Processing Employing Delta-E Metric and CNN- [14] In this article, it presents a point-of-care diagnostic system for the early diagnosis of Chronic Kidney Disease (CKD) based on images of test strips captured using mobile devices. The system applies image processing and Delta-E color

comparison in the LAB space to measure biomarker levels, in this case, albumin. To enhance reliability, a Convolutional Neural Network (CNN) is incorporated to advance preprocessing as well as minimize environmental noise. This hybrid strategy presents a precise, cost-effective, and accessible solution to the screening of CKD, particularly in low-resource environments.

Mohd Mohsin Ali, Janhavi Agrawal, Tripti Mishra and Manish Raj (2023):

An Automated Deep Learning Approach for Kidney Disease Detection- [15] A deep learning-based method for detecting and categorizing kidney diseases, such as kidney stones, tumors, and cysts, is presented in this study. Hyperparameter optimization and thorough validation were used to create and optimize a custom Convolutional Neural Network (CNN). The suggested model showed excellent precision and a validation accuracy of 99.62%. This work provides a dependable and effective way to improve the early diagnosis of kidney-related disorders by utilizing developments in image processing and predictive modeling.

Comparative Analysis

Sl.No	Authors	Techniques used	Results
1.	Narinder Kumar and Sanjay Singla (2024)[1]	Logistic Regression, Random Forest, Ensemble Methods	Highest performance from ensemble techniques; exact accuracy not stated but emphasized computational cost concerns
2.	Sonam Bhandurge, Dr. Kuldeep Sambrekar, Dr.Shridhar Allagi, Dr. Vijay S.Rajpurohit and Dr. Sanjeev Sannakki (2024)[2]	SVM, Random Forest, ANN, Logistic Regression	Random Forest achieved 100% accuracy on clinical data
3.	Garima Shukla, Gaurav Dhuriya, Sofia K Pillai and Aradhna Saini (2023)[3]	KNN, Decision Tree, ANN	Decision Tree showed highest accuracy at 98.60%
4.	Rama Al-Momani, Ghada Al-Mustafa, Razan Zeidan, Hiam Alquran, Wan Azani Mustafa and Ahmed Alkhayyat (2022)[4]	ANN, SVM, KNN	ANN achieved highest accuracy of 99.2%
5.	Pankaj Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawati, Tulika Chakrabarti, Michal Jasinski, Łukasz Jasinski and Radomir	ANN, C5.0, CHAID, Logistic Regression, L1 & L2 SVM, Random Tree, Deep Neural Network	DNN achieved highest accuracy of 99.6%; LSVM (L2) 98.86% with SMOTE & full features

	Gono (2021)[5]		
6.	M.S. Abirami, Sayak Das and Roopal Sood (2024)[6]	XGBoost, Random Forest	Achieved high accuracy (noted \~99.19%) using Boruta feature selection; emphasis on model efficiency
7.	Anurag, Narayan Vyas, Vishal Sharma and Deepak Balla (2023)[7]	Feature Selection + Classification (unspecified algorithms)	Achieved 89% accuracy; focused on adaptability and intelligent diagnostic support
8.	B.Seetha Lakshmi, Vijay Anand R, Habeeb Omotunde, M. Birundadevi, Mahalakshmi S and Madala Guru Brahmam (2024)[8]	Deep Learning with CNN for medical image classification	High accuracy, precision, and recall in classifying kidney lesions
9.	Akash Maurya, Rahul Wable, Rasika Shinde , Sebin John , Rahul Jadhav and Dakshayani.R (2019)[9]	ML classifiers (SVM, KNN, Decision Tree, Logistic Regression) + diet recommendation using blood potassium level	Accurate CKD stage classification + tailored diet plans
10.	Guozhen Chen, Chenguang Ding, Yang Li, Xiaojun Hu, Xiao Li, Li Ren, Xiaoming Ding, Puxun Tian, and Wujun Xue (2020)[10]	Adaptive Hybridized Deep CNN on IoMT platform	Effective early CKD prediction; improved accuracy via deep learning and IoMT integration
11.	Dupati Lohith Chowdari, Kesamreddy Vinay Kumar Reddy, Shoba L K and Pongiannan R K (2025)[11]	Multi-modal data fusion using CNN + MLP	Enhanced prediction accuracy through fusion of imaging and clinical data

12.	Vanathi.D, S.M. Ramesh, Sudha.K, Tamizharasu.K, Sengottaiyan.N and Kalyanasundaram.P (2024)[12]	KNN, SVM, ANN, and ensemble models like Random Forest, XGBoost	Achieved 99.2% accuracy; KNN and SVM performed better than ensemble methods
13.	Navaneeth Bhaskar and Suchetha M (2019)[13]	1D CNN + SVM classifier on salivary ammonia sensor data	Achieved 98.04% classification accuracy
14.	Espin Shalo S P, Jeeshan R J, S Gunasekar, Pius Agbulu G and Venkat Ramana N (2024)[14]	Delta-E Metric (LAB color space) + CNN for mobile image processing	Cost-effective and reliable early CKD detection in resource-limited settings
15.	Mohd Mohsin Ali, Janhavi Agrawal, Tripti Mishra and Manish Raj (2023)[15]	Custom 16-layer CNN (compared with MobileNet, DenseNet80, VisionNet)	99.62% validation accuracy; best performance among tested models

Results/Discussion

Machine learning (ML) has emerged as a useful asset for the early diagnosis of chronic kidney disease (CKD), facilitating effective diagnosis and treatment. High predicted accuracy has been observed in a variety of categorization models, including Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN), with RF maintaining 100% reliability in recognizing CKD. Additionally, ANN and SVM performed remarkably well, surpassing 98% accuracy, demonstrating ML's capacity to assist medical professionals in identifying CKD early on. Medical professionals can predict patient health and potentially reduce the need for costly operations like kidney transplants and dialysis by using machine learning (ML)-based diagnostics. Data pretreatment, which ensures that classification models receive clean and well-organized input, is a crucial component of applying ML-based CKD prediction.

Datasets are refined by eliminating unnecessary features and highlighting significant biomarkers like blood pressure, albumin levels, and serum creatinine using methods like Recursive Feature Elimination (RFE), Principal Component Analysis (PCA), and feature selection using deep learning. Missing values are handled using k-Nearest Neighbor (k-NN) imputation, which improves model accuracy and prevents biases.
Machine learning applications for patient population groupings become more predictable when cross-validation techniques are used to enhance model stability and reduce the likelihood of overfitting.
Despite these advancements, challenges remain before ML-based CKD prediction may be used in clinical settings.

Unbalanced data and variations in patient groups may limit the model's applicability to larger groups; further modifications would be required to guarantee accurate predictions in a range of groupings.
Furthermore, medical experts cannot predict the outcomes of most machine learning algorithms because they

are black boxes. Enhancing AI-powered diagnostics and promoting trust in them require explainable AI frameworks. To enhance ML-based CKD detection and its smooth integration into routine healthcare, more high-quality medical datasets must be available, and classification methods must be improved.

Conclusion

Chronic Kidney Disease (CKD) is a severe medical condition that advances subtly and goes unnoticed in most cases until it is advanced. Early diagnosis is essential for treatment and better patient outcomes. The survey stated above demonstrates how machine learning (ML) approaches can be applied to provide accurate, non-invasive, and reasonably priced methods of early CKD detection. Machine learning algorithms can detect subtle trends that are hard to spot with traditional diagnostic techniques by utilizing features from blood tests and CT scans. Feature selection is crucial for enhancing model performance since it only addresses the most important clinical parameters, saving computing time and preventing unnecessary medical procedures. Accuracy, sensitivity, and specificity are performance indicators that demonstrate the reliability of certain categorization algorithms, especially ensemble approaches. But the largest limitation in existing research is the utilization of small, unbalanced datasets that may compromise the models' generalizability. The future direction should emphasize increasing dataset size, enhancing class balance, and increasing model transparency to facilitate clinical uptake. In general, machine learning presents a promising strategy for early diagnosis of CKD, enabling quicker decision-making and lowering the dangers of delayed diagnosis and costly therapies such as dialysis or kidney transplantation.

References

[1] Narinder Kumar and Sanjay Singla, "Chronic kidney disease prediction: Optimization of machine learning algorithms" (2024)

[2] Sonam Bhandurge, Dr. Kuldeep Sambrekar, Dr.Shridhar Allagi, Dr. Vijay S.Rajpurohit and Dr. Sanjeev Sannakki, "A Proposed Model for Machine Learning-Based Early Chronic Kidney Disease Detection" (2024)

[3] Garima Shukla, Gaurav Dhuriya, Sofia K Pillai and Aradhna Saini, "Chronic Kidney Disease Prediction Using Machine Learning Algorithms and the Important Attributes for the Detection" (2023)

[4] Rama Al-Momani, Ghada Al-Mustafa, Razan Zeidan, Hiam Alquran, Wan Azani Mustafa and Ahmed Alkhayyat, "Chronic Kidney Disease Detection Using Machine Learning Technique" (2022)

[5] Pankaj Chittora, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawati, Tulika Chakrabarti, Michal Jasinski, Łukasz Jasinski and Radomir Gono, "Prediction of Chronic Kidney Disease - A Machine Learning Perspective." (2021)

[6] M.S. Abirami, Sayak Das and Roopal Sood, "Chronic Kidney Disease Identification using Random Forest and XGBoost." (2024)

[7] Anurag, Narayan Vyas, Vishal Sharma and Deepak Balla, "Chronic Kidney Disease Prediction Using Robust Approach in Machine Learning." (2023)

[8] B.Seetha Lakshmi, Vijay Anand R, Habeeb Omotunde, M. Birundadevi, Mahalakshmi S and Madala Guru Brahmam, "Deep Learning for Automated Classification of Kidney Lesions in Chronic Kidney Disease Patients." (2024)

[9] Akash Maurya, Rahul Wable, Rasika Shinde , Sebin John , Rahul Jadhav and Dakshayani.R, "Chronic Kidney Disease Prediction and Recommendation of Suitable Diet plan by using Machine Learning." (2019)

[10] Guozhen Chen, Chenguang Ding, Yang Li, Xiaojun Hu, Xiao Li, Li Ren, Xiaoming Ding, Puxun Tian, and Wujun Xue, "Prediction of Chronic Kidney Disease Using Adaptive Hybridized Deep Convolutional Neural Network on the Internet of Medical Things Platform." (2020)

[11] Dupati Lohith Chowdari, Kesamreddy Vinay Kumar Reddy, Shoba L K and Pongiannan R K, "Optimizing Chronic Kidney Disease Detection through Multi-Modal Data Fusion using CNN-MLP Architectures." (2025)

[12] Vanathi.D, S.M. Ramesh, Sudha.K, Tamizharasu.K, Sengottaiyan.N and Kalyanasundaram.P, "A Machine Learning Perspective for Predicting Chronic Kidney Disease." (2024)

[13] Navaneeth Bhaskar and Suchetha M, "A Deep-Learning-Based System for Automated Sensing of Chronic Kidney Disease." (2019)

[14] Espin Shalo S P, Jeeshan R J, S Gunasekar, Pius Agbulu G and Venkat Ramana N, "Automated CKD Detection via Image Processing Employing Delta-E Metric and CNN." (2024)

[15] Mohd Mohsin Ali, Janhavi Agrawal, Tripti Mishra and Manish Raj, "An Automated Deep Learning Approach for Kidney Disease Detection." (2023)

Figures and Tables



Figure1. Input Images

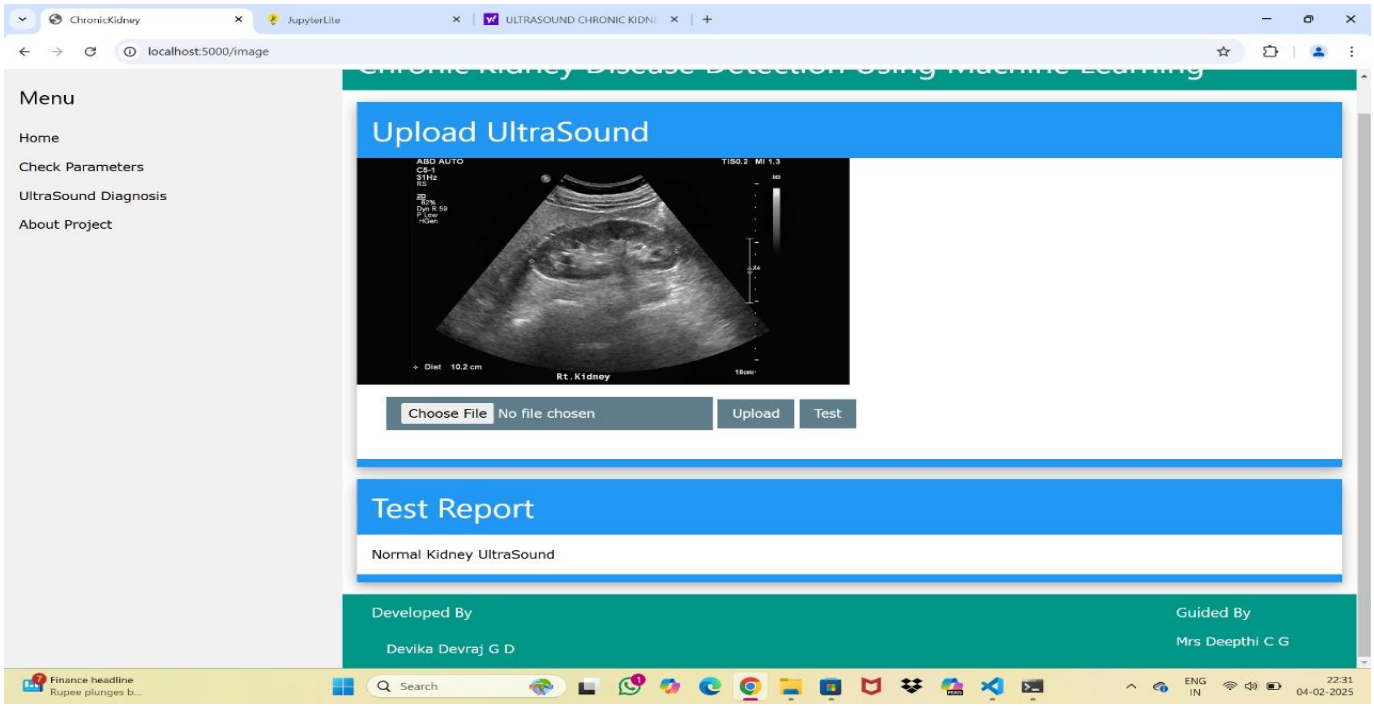


Figure2. Normal Kidney Test Report

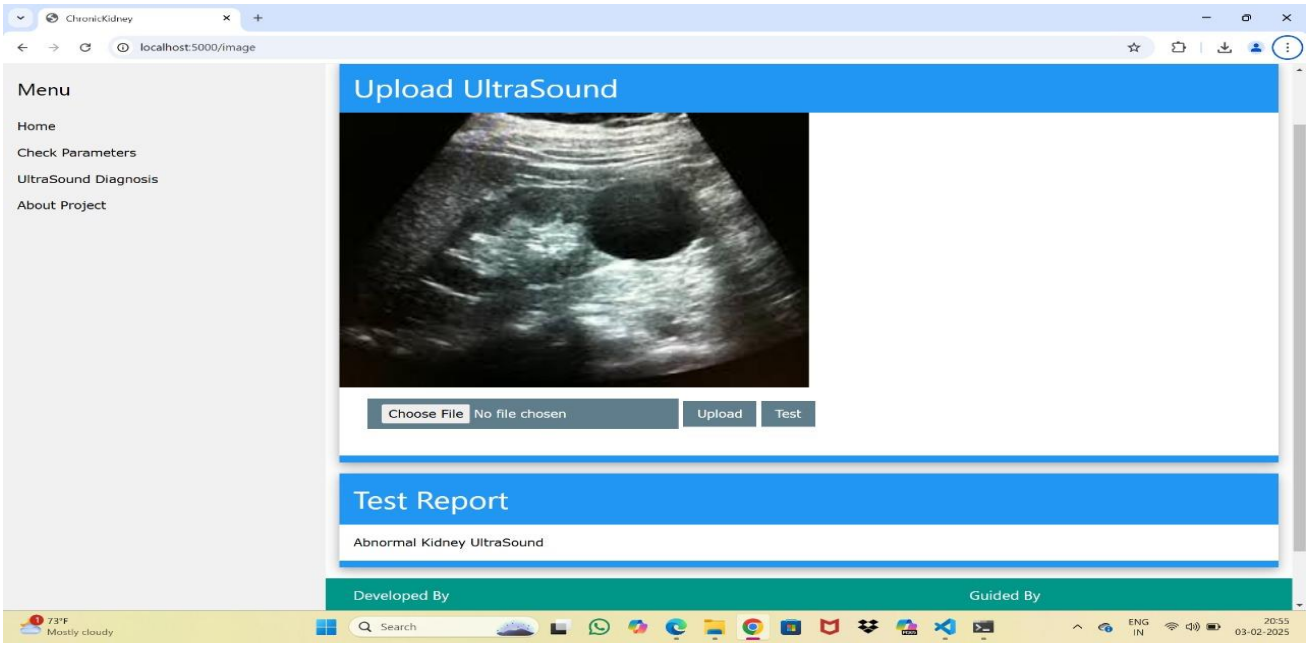


Figure3. Abnormal Kidney Test Report

Copyright & License:
© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.