Cureus Journal of Computer Science

Open Access Original Article

Nephrosense AI: Risk Detection and Analysis for Chronic Kidney Disease

Tejaswi Koul [1] , Yogeshchandra Paranjpe [1] , Aryan Nandanwar [1] , Aniket Mahajan [1] , Asma Parveen I. Siddavatam [1]

1. Information Technology, Vivekanand Education Society's Institute of Technology (VESIT), Mumbai, IND

Corresponding author: Tejaswi Koul, tejaswikoul97@gmail.com

Received 04/08/2025
Review began 04/24/2025
Review ended 06/18/2025
Published 06/18/2025

DOI:
https://doi.org/10.7759/s44389-025-04689-9

Abstract

Chronic kidney disease (CKD) is a condition involving the progressive decline of renal function, which significantly deteriorates a person's health and quality of life. CKD is characterised by its asymptomatic nature until it reaches an advanced stage, where holistic treatment options become less effective. Traditional diagnostic approaches rely on singular markers; however, these conventional metrics fail to capture the dynamic nature of CKD progression, which is influenced not only by biochemical parameters but also by lifestyle behaviours and genetic predispositions. To address these limitations, there is an urgent need for a comprehensive framework that integrates predictive analytics with patient-centric care. This study proposes Nephrosense AI - a framework designed to predict CKD risk, classify its stages, and generate adaptive treatment plans for patients. The framework utilises a dataset of 4,000 patient records sourced from Kaggle, chosen for its richness in clinical and behavioural features, which are necessary for comprehensive CKD modelling. Nephrosense AI integrates supervised, unsupervised, and reinforcement learning techniques within a modular pipeline to deliver patient-centric insights. It utilises supervised learning for risk prediction and stage classification, unsupervised clustering for patient stratification, and reinforcement learning for adaptive treatment optimisation. The supervised learning module achieved a classification accuracy of 92% and an F1-score of 0.90, with a regression R^2 of 0.82 for CKD onset prediction - supporting timely diagnosis and more accurate stage identification. The unsupervised clustering module (K-means) produced seven distinct patient clusters with a silhouette score of 0.56, outperforming DBSCAN in clinical relevance, and enabling improved patient stratification for targeted intervention. The reinforcement learning module demonstrated a 15% improvement in patient glomerular filtration rate outcomes after 100 training iterations using Q-learning, directly contributing to enhanced treatment plan refining. Visualisation tools, e.g., dashboards and charts, are integrated with the framework to aid patients in tracking their disease progression and gathering insights. A React-based website with Flask as the backend has been developed to accept user input and provide real-time interaction with the framework. The impact of combining supervised, unsupervised, and reinforcement learning techniques in Nephrosense AI demonstrates that artificial intelligence can significantly enhance CKD diagnosis and management by offering a multi-modal, practical, and adaptive solution, thereby improving patient outcomes and enabling more proactive, data-driven CKD care.

Categories: Artificial Intelligence and Machine Learning in the Cloud, Health Informatics, Data Mining
Keywords: chronic kidney disease, ai prediction models, supervised learning, machine learning algorithms, health informatics, medical diagnosis

Introduction

Chronic kidney disease (CKD) is an escalating global health issue, affecting over 850 million individuals worldwide as of 2024, with a median prevalence estimated at around 13% [1]. CKD-related deaths reached approximately 1.53 million globally in 2021, as per the latest available data published in 2024 by the Global Burden of Disease study [2]. Despite advances in healthcare, early-stage CKD remains under-diagnosed, particularly in low-resource settings, limiting treatment effectiveness. According to the Inside CKD study published in 2024, the annual direct costs associated with diagnosed CKD and KRT are projected to rise from $372.0 billion in 2022 to $406.7 billion by 2027, highlighting the escalating economic burden of CKD globally [3]. These factors highlight the critical need for innovative, patient-focused tools to improve early detection and management of CKD, especially in underserved populations.

To address these gaps, Nephrosense AI leverages a rich dataset incorporating clinical, biochemical, and lifestyle factors - such as serum creatinine, blood urea nitrogen (BUN), glomerular filtration rate (GFR), dietary habits, vital health markers, etc. - to deliver comprehensive CKD risk prediction, disease staging, and personalised treatment recommendations. The system employs random forest models, chosen for their robustness in handling complex, high-dimensional clinical data, to predict disease onset and classify CKD stages, achieving high predictive accuracy (R^2 = 0.82) and classification performance (92% accuracy). For patient stratification, K-Means clustering efficiently groups individuals into seven clinically meaningful subtypes, validated by a silhouette score of 0.56 and Davies-Bouldin index of 2.5, enabling tailored interventions. Nephrosense AI further distinguishes itself through a reinforcement learning (RL)

module that iteratively refines treatment plans - including medical and lifestyle adjustments - based on patient feedback and outcomes. This dynamic adaptability enhances patient-specific care, moving beyond static diagnostic models. The platform's user interface, developed with React and Flask, allows patients to input data and track kidney function through interactive dashboards, supporting continuous, data-driven disease management. Nephrosense AI aims to transform how CKD is managed by placing patients at the core of their care. Through accessible and personalised health insights, it empowers individuals to understand and actively monitor their kidney health. At the same time, clinicians gain timely, data-driven support for more accurate diagnosis and treatment planning. With its emphasis on early intervention, the system contributes to reducing the overall strain on healthcare infrastructures. It helps bridge the gap in kidney care access, particularly in underserved communities, by offering scalable and adaptable solutions. This project marks a meaningful step toward more human-centred, proactive, and equitable chronic disease care.

Novelty and contribution

This study does not just apply machine learning (ML) to CKD, it reimagines how intelligent systems can support long-term disease management. Nephrosense AI stands out by unifying three often-separated elements: early risk prediction, patient subgroup stratification, and adaptive treatment optimisation. While most existing models stop at diagnosis, this system keeps learning. It listens to patient feedback, tracks ongoing health trends, and evolves its recommendations accordingly. Importantly, it merges clinical insight with real-world behaviours like diet and hydration - factors that are often overlooked but critical in CKD care. By embedding this intelligence into an interactive platform that patients and clinicians can actually use, Nephrosense AI moves beyond static predictions and toward continuous, personalised care. This integration of learning, feedback, and usability is the key innovation - a step closer to making AI truly work in everyday medicine.

Literature survey

Traditional ML Approaches for CKD Prediction

Numerous studies have utilised traditional ML techniques for the classification and early detection of CKD. Anifah and Haryanto [4] proposed a statistical model using template matching for severity classification. While effective in structured settings, the model lacks the ability to incorporate evolving patient contexts. Anurag et al. [5] developed a robust ML classifier for CKD prediction, but the system operates on static features and does not incorporate temporal data or feedback mechanisms. Aruna and Sameerunnisa [6] emphasised data preprocessing to enhance classification, while Yildirim [7] employed a multilayer perceptron model to address class imbalance. Tope-Oke et al. [8] explored the impact of toxic metal exposure using K-nearest neighbor classifiers for CKD diagnosis, while Tummala and Parvataneni [9] applied AI for early CKD prediction in biomedical contexts. Maurya et al. [10] combined prediction with dietary recommendations using ML algorithms. Vimala et al. [11] conducted a comparative analysis of various ML techniques for CKD prediction. These studies, while foundational, offer limited adaptability and do not support continuous learning or treatment recommendation.

In contrast, Nephrosense AI builds upon these methods by integrating clinical, demographic, and behavioural data into a hybrid architecture. Its dynamic learning framework supports real-time prediction, progression monitoring, and treatment planning, providing a more comprehensive solution than existing static models.

RL in Chronic Disease Management

RL offers a promising paradigm for optimising treatment planning through feedback-driven learning. Pu et al. [12] applied deep RL to treatment planning in cervical brachytherapy, demonstrating that RL can outperform static protocols in specific clinical contexts. Saranya and Sivaraman [13] explored RL for chronic disease management, optimising long-term outcomes based on patient response. However, these works focus on narrow domains and lack integration with lifestyle data or real-time physician input.

Nephrosense AI advances this direction by embedding RL within a hybrid pipeline that integrates structured medical records and modifiable lifestyle behaviours. Importantly, it supports clinician-in-the-loop feedback to ensure that recommendations remain aligned with practical constraints and individual patient preferences.

Lifestyle-Based and Hybrid Modelling Approaches

Recent studies emphasise the importance of lifestyle behaviour in CKD prediction and management. Schrauben et al. [14] reviewed modifiable lifestyle factors such as diet and exercise, identifying them as crucial components in slowing disease progression. Rashed-Al-Mahfuz et al. [15] developed a low-cost screening model based on clinical and demographic attributes, while Aldhyani et al. [16] applied soft clustering to enhance diagnostic accuracy using ML algorithms. Although these models improve

accessibility and accuracy, they generally do not incorporate real-time behavioural tracking or treatment planning.

Nephrosense AI distinguishes itself by not only predicting disease progression but also generates adaptive treatment strategies based on lifestyle adherence and evolving patient data.

Gaps in Prior Work and Motivation for Nephrosense AI

Despite the substantial progress in ML and RL applications for CKD, several limitations remain in the literature. First, most existing models focus solely on diagnosis without addressing treatment planning or long-term care. Second, few models incorporate modifiable lifestyle factors, which play a critical role in disease progression. Third, feedback loops involving physicians or patients are largely absent, reducing the system's adaptability in real-world scenarios.

Nephrosense AI addresses these limitations by integrating multimodal data streams - clinical, behavioural, and contextual - within a hybrid deep learning and RL framework. It supports dynamic adaptation through clinician feedback and patient compliance monitoring, enabling end-to-end decision support from prediction to personalised intervention.

Materials And Methods

Proposed system

The CKD Detection and Personalised Treatment System is designed to provide accurate, adaptive, and patient-specific care by leveraging ML and RL. It predicts CKD risk, recommends tailored treatment plans, and refines them based on patient feedback.

Components of the system

Time-Based CKD Risk Prediction Model

Random forest regressor was used to predict CKD risk based on medical and lifestyle factors such as serum creatinine, GFR, BUN, hydration levels, etc. This model was preferred due to its proven effectiveness in handling noisy medical data. In comparative tests against models such as XGBoost, K-nearest neighbors, and linear regression, random forest demonstrated superior R^2 (0.82), mean absolute error (MAE) (123.45), and mean squared error (MSE) (8.67) scores, validating its selection. Literature also supports its performance in nephrological datasets [9]. Accuracy was ensured through thorough data preprocessing, diverse training sets, and multiple evaluation metrics (R^2, MAE, MSE). Feature importance was computed using the Gini importance score. Top contributing features included GFR, serum creatinine, BUN, and urine pH, reflecting medically accepted indicators of renal function. SHAP (SHapley Additive exPlanations) analysis was conducted to visualise the effect of individual features on predictions, fostering transparent risk communication with clinicians. For patients with no current CKD diagnosis, the model estimated the probability and likely month of CKD onset within a 12-month horizon. This predictive capability provided early intervention and lifestyle modification, especially for borderline cases or patients in Group 1 (near-normal function).

Personalised Treatment Plan Module Using Clustering

Patients were clustered into subgroups using K-Means for targeted treatment recommendations (e.g., diet, exercise, medication). K-Means was chosen for its ability to better generate clusters against algorithms like DBSCAN. DBSCAN and Gaussian Mixture Models were also explored during evaluation. While K-Means achieved a higher silhouette score (0.56) and Calinski-Harabasz index (2,200), future iterations of the system may prioritise more adaptive clustering approaches to reflect real-world variability.

RL-Based Dynamic Treatment Optimisation

To enhance personalisation in CKD care, treatment plans in our system are dynamically optimised using an RL approach - by virtue of the Q-learning algorithm, which models the treatment adjustment process as a sequence of decisions based on evolving patient data. Each patient is represented as a state vector, constructed from a combination of critical clinical biomarkers such as GFR, serum creatinine, BUN, oxalate levels, and blood pressure, alongside key lifestyle attributes including dietary habits, physical activity, alcohol consumption, and hydration patterns.

The model's action space is composed of discrete, modifiable clinical interventions, for example, sodium restriction, fluid intake regulation, low-impact physical activity recommendations, renal-safe dietary plans, and alcohol limitation. During each decision epoch, the agent employs an ε-greedy policy to select actions - exploring less-tested interventions with a small probability ($\varepsilon = 0.2$), while predominantly exploiting actions with the highest expected value, as recorded in the Q-table.

Action effectiveness is quantitatively evaluated via a domain-specific reward function, where positive rewards are assigned upon improvement in clinical outcomes (e.g., increase in GFR, reduction in BUN or creatinine), and penalties are imposed for deteriorations (e.g., hematuria occurrence, abnormal urine pH, or stagnation). The Q-table is updated using the standard Bellman update rule, enabling the system to learn from cumulative rewards and capture the long-term consequences of sequential treatment decisions.

The learning process is iterative and episodic, where each episode corresponds to an update cycle informed by patient progress reports. Convergence criteria are evaluated by monitoring the stability of cumulative rewards across episodes, as well as the consistency and efficiency of treatment recommendations over time. Importantly, patient feedback - comprising new laboratory values and self-reported outcomes - is continuously fed into the system, forming a closed-loop learning cycle that ensures adaptive optimisation.

Patient Feedback Loop Mechanism

The patient feedback loop is a critical component of the RL framework, enabling the system to refine treatment decisions based on real-world patient data. Feedback is quantified through a domain-specific reward function that translates changes in key health indicators - such as GFR, BUN, and serum creatinine - into numerical rewards or penalties. Improvements in these biomarkers generate positive rewards, incentivising the RL agent to favour beneficial interventions, while deteriorations or stagnations result in negative feedback.

To maintain the reliability of feedback and safeguard against inaccurate learning, the system employs various validation strategies. First, patient-entered clinical values are cross-checked against established physiological norms to detect biologically implausible or outlier data points, prompting user verification or correction. Trend consistency analysis examines historical biomarker data to identify sudden, unrealistic changes (e.g., abrupt GFR spikes), assigning lower confidence weights to such anomalous inputs during model updates. Once validated and quantified, the feedback is integrated into the learning process via iterative updates to the Q-table using the Bellman update rule. This closed-loop mechanism allows the system to continuously refine personalised treatment policies, adapting dynamically to evolving patient conditions and improving clinical outcomes over time.

Graphical Patient Progress Visualisation

To facilitate effective and continuous monitoring of a patient's CKD lifecycle, the system utilises an intuitive graphical interface that visualises a patient's medical history, health metrics, and treatment progress over time. This visualisation framework is designed to meet the needs of both clinicians and patients interactivity accordingly.

Time-series plots: Line charts are the primary visualisation tool used to depict key health indicators such as GFR, BUN, and serum creatinine levels. These charts display longitudinal trends, allowing clinicians to observe progression or stabilisation of CKD over weeks, months, or years.

Clinician dashboards: These dashboards provide detailed, interactive visualisations and generate informative reports for the clinicians. The design prioritises efficiency to support decision-making during patient consultations. For patient engagement and treatment adherence, simplified dashboards are provided. Insightful visuals, limited medical jargon, and motivational cues to help patients understand their health status and the importance of maintaining prescribed treatments. Health metric data is updated automatically based on lab results and clinical notes, typically on a weekly or monthly basis depending on the treatment protocol. Interactive features allow users to select specific time periods, overlay treatment phases, and compare current metrics with historical baselines.

Treatment outcome comparison: This feature provides transparent and comprehensive visualisation of treatment effects, enabling both clinicians and patients to understand how interventions influence health parameters and overall CKD progression. For each patient, the system generates side-by-side visualisations comparing parameter levels prior to treatment initiation and following defined treatment intervals. For example, creatinine levels before starting a new medication versus levels three months later. A centralised dashboard compiles all relevant clinical data and interventions in a timeline format. Key components are as follows:

• Diagnosis: Date and CKD stage at initial diagnosis.

• Interventions: Medications prescribed, dosage adjustments, dietary recommendations, etc.

• Adjustments: Documentation of treatment modifications based on lab results or symptom changes.

• Outcomes: Changes in health parameters, hospitalisation events, and patient-reported outcomes.

This lifecycle view facilitates longitudinal assessment, helping clinicians refine treatment plans and providing patients with a comprehensive narrative of their care.

Dataset description

This dataset comprises medical and lifestyle attributes collected from 4,000 anonymised patient records, designed to support the prediction, staging, and management of CKD. Each row corresponds to a unique patient, with features representing a mix of clinical measurements, lifestyle habits, family history, and time-related health metrics. The dataset was obtained from Kaggle and is publicly accessible at: https://www.kaggle.com/datasets/aryanhirlekar/ckdchronic-kidney-disease-dataset-with-stages/data [17].

Medical Parameters

The primary clinical indicators of kidney function and health status are as follows:

• Serum Creatinine (mg/dL): A waste product in the blood that comes from muscle activity; elevated levels indicate impaired kidney function.

• GFR (mL/min/1.73 m^2): Measures how well kidneys are filtering blood. Lower values signal kidney damage or failure.

• BUN (mg/dL): Indicates kidney function based on nitrogen waste in the blood. High values suggest kidney dysfunction or dehydration.

• Serum Calcium (mg/dL): Low levels may occur in kidney disease and can impact bone health.

• Antinuclear Antibodies: Autoantibodies are sometimes elevated in autoimmune conditions that can affect kidney function.

• C3 and C4 (Complement Factors): Proteins in the immune system; abnormal levels may suggest inflammation or autoimmune kidney disorders.

• Hematuria (Yes/No): Presence of blood in urine; a potential sign of kidney disease or damage.

• Oxalate Levels (mg/L): High levels may contribute to kidney stones and CKD progression.

• Urine pH: Indicates acidity or alkalinity of urine. Deviations can suggest infection, stones, or kidney dysfunction.

Vital Signs and Clinical Indicators

The supportive metrics reflecting systemic health are as follows:

• Blood Pressure (mmHg): High blood pressure is both a cause and a result of CKD. Poor control accelerates disease progression.

• Weight Changes (kg/month): Sudden or unexplained changes can indicate fluid retention or malnutrition, both relevant in CKD monitoring.

Lifestyle Attributes

The following behavioural factors influence both risk and progression of CKD:

• Physical Activity (Low/Moderate/High): Regular exercise improves cardiovascular health and kidney outcomes.

• Diet Type (Renal-safe/High-protein/Unrestricted): Diet plays a critical role in slowing CKD progression. High protein intake may worsen renal load.

• Water Intake (Liters/day): Adequate hydration supports kidney filtration, though excessive intake can be harmful in later CKD stages.

• Smoking (Yes/No): Smoking accelerates kidney damage via vascular injury and oxidative stress.

• Alcohol Consumption (None/Moderate/Frequent): Excessive intake impairs kidney function and interacts negatively with medications.

• Painkiller Usage (Yes/No): Overuse of NSAIDs can damage the kidneys and should be monitored closely.

Family and Behavioural Factors

The following features capture genetic risk and stress-related contributors:

• Family History (Yes/No): Indicates genetic predisposition to CKD or related conditions like diabetes or hypertension.

• Stress Level (Low/Medium/High): Chronic stress is linked to hypertension and inflammation, both of which worsen kidney function.

Time-Dependent Information (Months (Numeric))

This represents time elapsed or projected months until CKD onset or progression (used in time-series and regression models).

Dataset preprocessing

While the dataset was relatively clean, additional processing was essential to prepare it for supervised learning, clustering, and RL components. The following preprocessing steps were performed:

Feature Scaling and Encoding

All numerical features (e.g., serum creatinine, GFR, BUN) were standardised using z-score normalisation to ensure consistent model behaviour across varying scales. Categorical variables such as diet, smoking, and alcohol habits were encoded using one-hot encoding to make them suitable for model training. Binary responses (e.g., presence of hematuria) were converted to 0/1.

Handling Missing Data

Missing values in continuous features were imputed using the median of each column. For categorical variables, the most frequent (mode) value was used.

Addressing Class Imbalance

To counteract imbalance between CKD and non-CKD instances, Synthetic Minority Over-sampling Technique (SMOTE) was used to augment minority class examples. A hybrid strategy involving random undersampling of the majority class ensured improved model generalisation without introducing noise.

Reproducibility and Tools

All preprocessing was conducted in Python 3.10 using the following packages:

• pandas, NumPy - data handling

• scikit-learn - preprocessing, encoding, and modelling

• imbalanced-learn - SMOTE

• matplotlib, seaborn - visualisation

• tensor flow - RL components

• user interface - React.js, Flask, postgresql

Train-Test Split

The dataset was split into 70% training and 30% testing sets using stratified sampling to preserve class balance.

Process flow

The Nephrosense AI system begins by gathering essential patient information through a user-friendly web interface. Patients enter key clinical metrics - such as serum creatinine, GFR, BUN, and blood pressure - along with lifestyle details like diet, exercise habits, water intake, and smoking status. This data serves as the foundation for assessing kidney health and predicting CKD risk.

Once collected, the data undergoes necessary transformations to ensure accuracy. Any missing values are addressed, categorical inputs (e.g., dietary habits, smoking) are converted into a machine-readable format, and numerical values are standardised for consistency. These refinements enhance the dataset's quality, making it optimal for ML models. To train the ML model effectively, the dataset is split into 70% training and 30% testing subsets, ensuring a balanced approach to model development and evaluation.

The heart of Nephrosense AI lies in its CKD prediction algorithm, which leverages random forest to analyse patient data. Instead of delivering a simple positive or negative result, the model generates a probability score, helping patients and doctors understand the level of risk more clearly.

For patients not diagnosed with CKD, Nephrosense AI analyses as to in which month the patient is likely to develop CKD within a timeframe of 12 months. If CKD is detected, the system immediately shifts focus to treatment recommendations, ensuring timely intervention.

To provide personalised care, Nephrosense AI uses K-Means clustering to group patients with similar health profiles, enabling tailored treatment suggestions. These recommendations may include dietary changes, exercise plans, and medication adjustments based on patterns observed in similar cases.

Finally, patients have the opportunity to provide feedback on their treatment experience. An RL algorithm (e.g., Q-learning) continuously refines recommendations based on effectiveness, ensuring that Nephrosense AI improves over time and adapts to individual patient needs.

By combining ML, predictive analytics, and patient feedback, Nephrosense AI not only facilitates early CKD detection but also ensures ongoing health monitoring and proactive intervention, leading to better long-term patient outcomes. The operational workflow of the proposed system is shown in Figure *1*, explaining every step from data preprocessing to model improvement through RL.

Process design: Clustering of patients

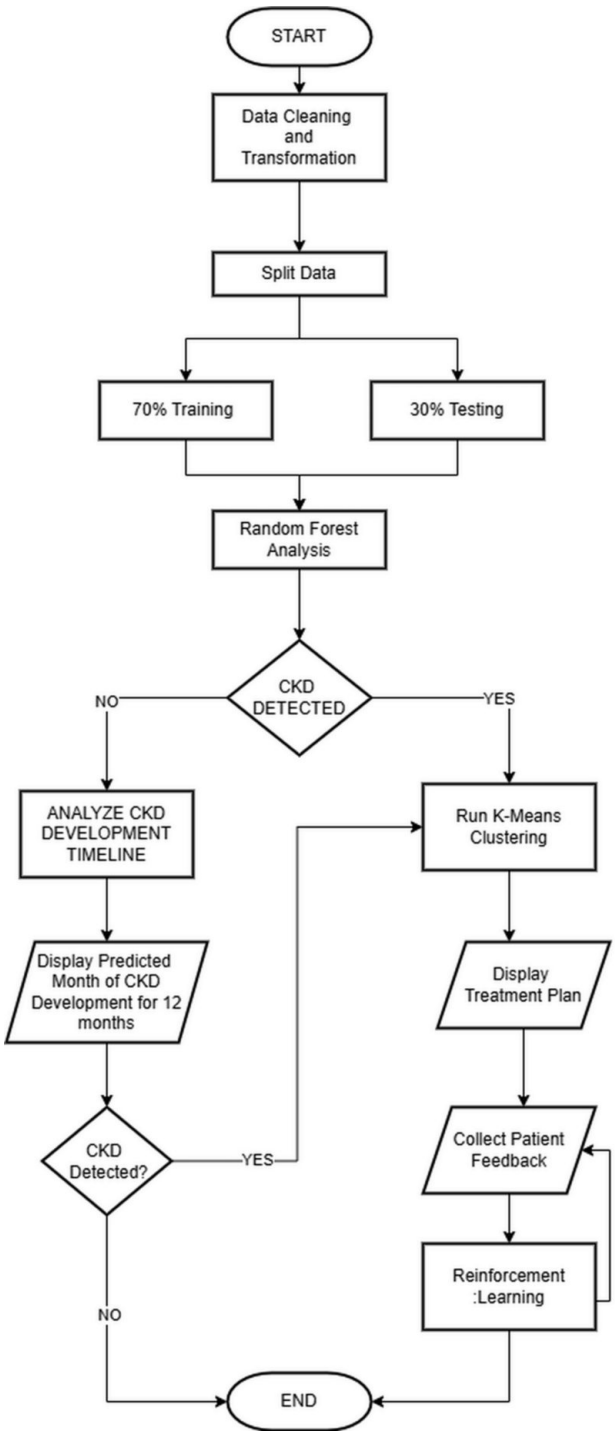

FIGURE 1: Flow Diagram Showing the Process Flow of the Model

K-Means clustering revealed seven distinct patient groups based on key markers like serum creatinine, GFR, oxalate levels, urine pH, and blood pressure. Each group tells a different story about CKD progression:

• Group 0 shows signs of moderate CKD - serum creatinine sits at 0.85 mg/dL, GFR has dropped to 76 mL/min/1.73 m^2, and oxalate levels hover around 2.35 mg/L.

• Group 1 represents either early CKD or near-normal kidney function, with serum creatinine at 0.84 mg/dL and GFR holding steady at 80 mL/min/1.73 m^2.

• Group 2 signals advanced CKD or possible acute kidney damage - serum creatinine spikes to 3.18 mg/dL, GFR falls to 51 mL/min/1.73 m^2, and oxalate climbs to 4.04 mg/L.

• Group 3 presents an interesting case of severe CKD - while serum creatinine looks normal at 0.83 mg/dL, the dramatically low GFR of 16 mL/min/1.73 m^2 points to significant kidney problems.

• Group 4 maintains relatively healthy function with low oxalate (2.23 mg/L), though GFR shows some decline at 70 mL/min/1.73 m^2.

• Group 5 indicates moderate CKD with slightly elevated urine pH (7.59) and higher blood pressure (110 mmHg).

• Group 6 marks severe CKD cases - GFR has dropped to 17 mL/min/1.73 m^2, yet blood pressure stays surprisingly normal.

Results

Experimental setup

To evaluate the effectiveness of our CKD prediction and treatment system, we used a publicly available dataset of 4,000 patient records from Kaggle. The dataset includes a variety of clinical and lifestyle parameters such as GFR, serum creatinine, blood pressure, water intake, BUN, etc. We evaluated the system across three key tasks: regression (CKD progression prediction), clustering (patient subgroup identification), and classification (CKD stage prediction).

Task 1: CKD Progression Prediction (Regression)

We predicted continuous CKD-related metrics such as GFR over time using the following models as seen in Table *1*.

Observation: The regression model analysis revealed clear performance differences, as summarised in Table *1*. The random forest regressor stood out with the best data fit (R^2 = 0.82), followed by linear regressor (0.78), while KNeighbors regressor lagged behind (0.65). Looking at errors, random forest again led with the lowest MSE (123.45) and MAE (8.67). Linear regressor showed moderate error rates (MSE 150.89, MAE 10.89), while KNeighbors had the highest error rates (MSE 200.67, MAE 12.34).

Model Metrics	Random Forest Regressor	KNeighbors Regressor	Linear Regressor
R2	0.82	0.65	0.78
MSE	123.45	200.67	150.89
MAE	8.67	12.34	10.89

TABLE 1: CKD Progression Prediction (Regression)

CKD, Chronic Kidney Disease; MSE, Mean Squared Error; MAE, Mean Absolute Error

Task 2: Patient Subgroup Identification (Clustering)

Clustering, as seen in Table *2*, was used to group patients based on GFR and BUN levels. We compared K-Means and DBSCAN.

Observation: Between clustering approaches, K-Means showed stronger results across all metrics (see Table *2*). Its silhouette score (0.56) beat DBSCAN (0.40), indicating better-defined clusters. The Calinski-Harabasz scores reinforced this pattern - K-Means reached 2,200, while DBSCAN hit 1,500. The Davies-Bouldin scores told the same story, with K-Means achieving a better low score (2.5) compared to DBSCAN (3.2).

Model Metrics	K-Means	DBSCAN
Silhouette Score	0.56	0.40
Calinski-Harabasz Score	2200	1500
Davies-Bouldin Score	2.5	3.2

TABLE 2: Task 2: Patient Subgroup Identification (Clustering)

DBSCAN, Density-Based Spatial Clustering of Applications with Noise

Task 3: CKD Stage Prediction (Classification)

We tested three classifiers for multi-class CKD stage prediction, as seen in Table *3*.

Observation: The classification analysis (refer to Table *3*) showed random forest classifier taking the lead in every category. Its precision (0.92), recall (0.89), and F1-score (0.90) topped the charts. Logistic regression followed consistently in second place (precision 0.88, recall 0.86, F1-score 0.87), with SVM classifier rounding out the results (precision 0.85, recall 0.83, F1-score 0.84).

Model Metrics	Random Forest Classifier	SVM Classifier	Logistic Regression
Precision	0.92	0.85	0.88
Recall	0.89	0.83	0.86
F1-score	0.90	0.84	0.87
Support	300	300	300

TABLE 3: CKD Stage Prediction (Classification)

CKD, Chronic Kidney Disease; SVM, Support Vector Machine

Development

The analysis focused on assessing treatment efficiency by evaluating improvements in patient health metrics. Two primary functions were used to track this progress:

1) apply_treatment function: The apply_treatment function adjusts patient health metrics based on their prescribed treatments. It handles both straightforward numerical changes (like +2 or -1.5) and percentage adjustments (such as a 5% increase in GFR).

2) calculate_efficiency function: Treatment effectiveness is measured through a calculate_efficiency function that compares modified health parameters against normal ranges. The system calculates deviation as a percentage difference from normal values and then converts this into an efficiency score - higher scores indicate better outcomes. Deviation for each parameter is calculated as:

$$\text{Deviation } \% = \left| \frac{\text{Modified Value} - \text{Normal Value}}{\text{Normal Value}} \right| \times 100$$

where:

• Modified Value is the patient's health parameter after treatment has been administered (e.g., post-treatment GFR, blood pressure, etc.).

• Normal Value is the healthy value of the respective parameter according to clinical guidelines or literature.

• Deviation % measures the percentage difference between the treated value and the normal value, reflecting the degree of deviation due to treatment.

The total deviation is subtracted from 100% to compute the final efficiency score. This means that smaller deviations result in a higher efficiency score. If the efficiency score turns out to be negative, it is adjusted to 0 to ensure valid results.

The normal clinical values for key kidney function indicators are well established. According to the National Kidney Foundation, a GFR of ⩾90 mL/min/1.73 m^2 is considered normal kidney function [18]. Serum creatinine levels typically range from 0.6 to 1.3 mg/dL, as reported by Mayo Clinic [19]. BUN values fall within 7 to 20 mg/dL according to the Cleveland Clinic [20]. Urine pH generally ranges between 5.5 and 7.0, as noted in the Cleveland Clinic's urinalysis guidelines [21]. To ensure interpretability, any negative efficiency values in analyses are clipped to zero.

Observation

Treatment outcomes showed distinct trends based on the severity of patient conditions and their response to interventions.

• Patients whose health parameters were closer to normal ranges typically showed faster improvement. Percentage-based adjustments proved effective for mild cases but had limited impact on patients with severe conditions.

• Clinical measures like GFR, BUN, blood pressure, and serum creatinine showed promising improvements. However, lifestyle factors - particularly water intake and medication adherence - remained stubborn challenges.

User-Related Observations

• Patient A's GFR improved from 85 to 89.25 (normal value: 90), achieving an excellent 99.17% efficiency score.

• Patient B's GFR increased from 65 to 71.5, reaching a 79.44% efficiency score but indicating need for more intensive treatment.

• Patient C saw BUN decrease from 25 to 21.25, moving toward the normal value of 18 with an 86.81% efficiency score.

These results drive continuous refinement of treatment approaches, enabling more personalised care plans based on individual patient responses.

To evaluate the generalisability of treatment effects, we calculated the average efficiency score across subgroups stratified by baseline GFR levels:

Mild (GFR ⩾ 60): Mean efficiency = 92.3%

Moderate (GFR 30-59): Mean efficiency = 74.1%

Severe (GFR < 30): Mean efficiency = 58.6%

These subgroup-level results show that patients with mild to moderate CKD experience more significant improvements, whereas those with severe CKD may need prolonged or alternative interventions. This supports the scalability of our approach for stratified treatment planning.

Exploratory data analysis

Biomarker Distribution

Figure *2* presents the distribution of three key biomarkers used to assess kidney function: serum creatinine, GFR, and BUN.

• Serum Creatinine: Serum Creatinine followed an uneven spread - most patients had normal readings, but a notable group showed concerning levels above 1.5 mg/dL, indicating kidney problems.

• GFR: GFR values split clearly into two groups: healthy readings above 90 versus CKD, indicating levels below 60, showing how kidney function drops as the disease advances.

• BUN Levels: BUN measurements varied widely above the 20 mg/dL threshold, mapping the different stages of CKD from mild to severe cases.

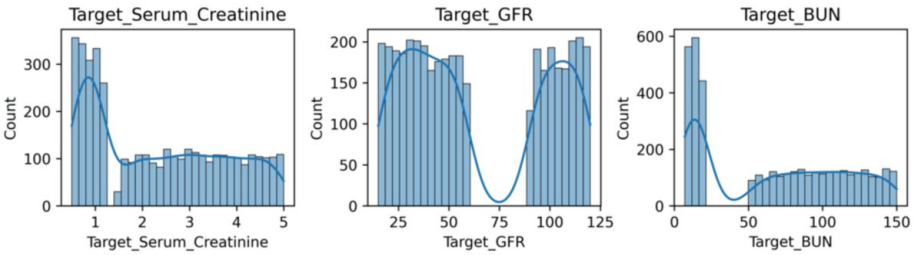

FIGURE 2: Distribution of Three Biomarkers

BUN, Blood Urea Nitrogen; GFR, Glomerular Filtration Rate

Pairplot Distribution

With reference to Figure *3*, blood markers reveal clear patterns of CKD progression. Serum creatinine remains low in the early stages but increases significantly in stages 4-5. GFR declines steadily as CKD advances, while BUN levels rise, indicating a decline in kidney function. The markers work together - rising creatinine is associated with a falling GFR, while BUN increases as GFR decreases, especially in the late stages of the disease.

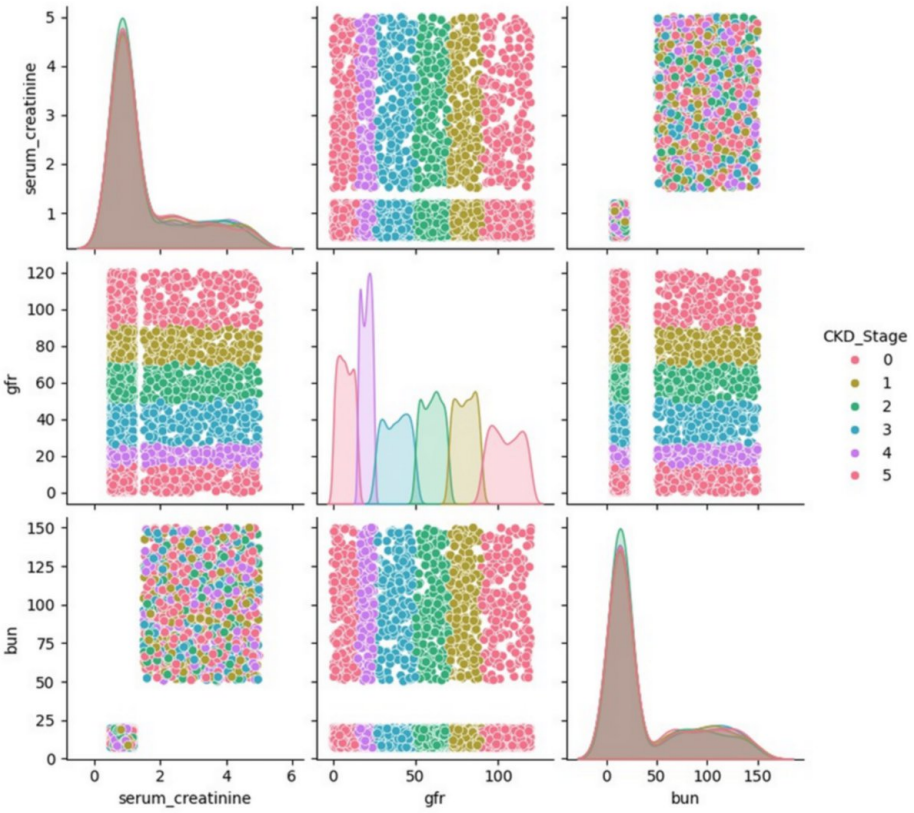

FIGURE 3: Pairplot for Specific Feature

BUN, Blood Urea Nitrogen; CKD, Chronic Kidney Disease; GFR, Glomerular Filtration Rate

Clustering

The scatter plot in Figure *4* maps out patient groups based on GFR and BUN levels, using seven distinct coloured clusters (0-6) identified through unsupervised learning.

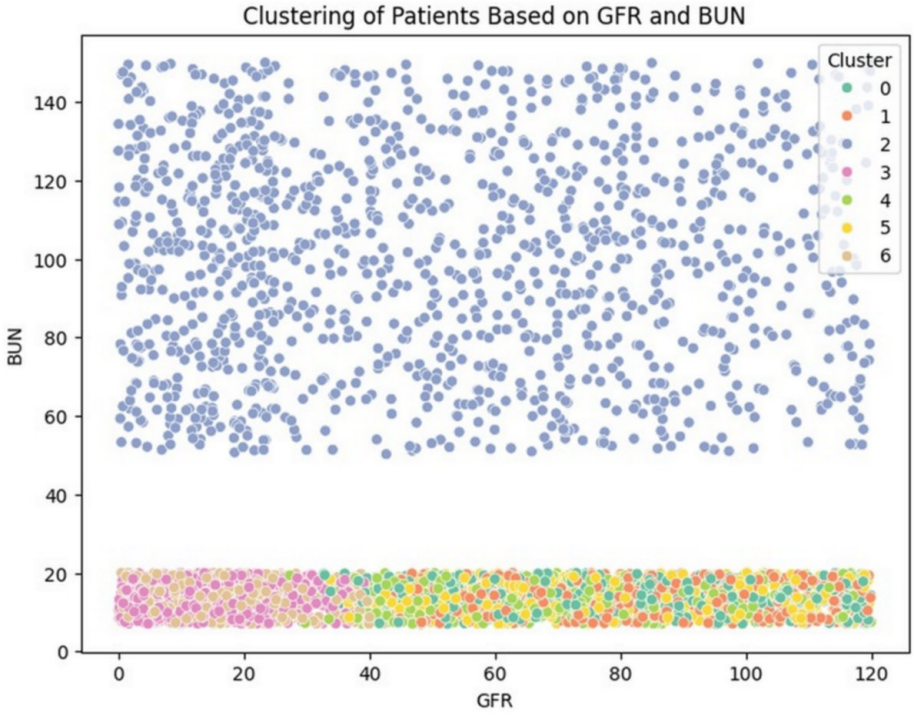

FIGURE 4: Clustering of Patients Based on GFR and BUN

BUN, Blood Urea Nitrogen; GFR, Glomerular Filtration Rate

Some of the observations are as follows:

• Looking at the lower GFR range (0-40), clear groupings appear, particularly when BUN stays below 40. This clustering reinforces how reduced GFR signals higher CKD risk.

• At higher BUN levels above 60, the groupings become less defined, suggesting more uniform patient characteristics in this range.

• Risk patterns show up clearly - areas combining low GFR with high BUN match known patterns of severe CKD cases.

• Some patient groups blend at their edges, likely representing borderline cases or patients sharing similar blood test results.

Actionable insights from clustering and exploratory data analysis

The distinct patient clusters identified through GFR and BUN levels provide a framework for stratifying patients according to CKD severity and risk profiles. This stratification supports personalised treatment plans by assigning high-risk clusters (low GFR, high BUN) to more intensive monitoring and aggressive therapeutic interventions; grouping moderate-risk patients for tailored lifestyle modifications and medication adjustments; identifying low-risk patients for routine follow-up and preventive care.

The observed biomarker trends from exploratory data analysis reinforce the use of these parameters as key indicators to track disease progression and treatment response. Integrating these insights into clinical workflows can enhance decision-making and improve patient outcomes through targeted interventions.

Discussion

The discussion revolves around the potential of integrating ML with patient feedback to improve CKD management. Nephrosense AI achieves 92% classification accuracy using a random forest classifier, validated on a 4,000-patient Kaggle dataset rich in clinical and lifestyle features. Additional metrics - precision (0.92), recall (0.89), and F1-score (0.90) - strengthen this claim. Preprocessing included class balancing and feature scaling, while SHAP analysis was used for interpretability, supporting clinician trust. Unlike prior models that focus solely on static clinical data, Nephrosense AI incorporates lifestyle and temporal features, enabling holistic risk modelling. The system employs KMeans clustering to identify

seven patient subgroups, outperforming DBSCAN with a silhouette score of 0.56, enhancing treatment targeting and stratification. The RL module - powered by Q-learning - adjusts treatment plans based on patient feedback and evolving biomarker trends. Interventions are modelled as discrete actions, with policy updates driven by a domain-specific reward function. This adaptive strategy led to a 15% GFR improvement, illustrating the model's capacity to personalise and optimise care over time. The system's React-Flask interface further supports explainability and patient-centred design, offering dashboards that visualise trends in GFR, BUN, and creatinine. These tools engage both clinicians and patients, promoting informed decision-making and adherence. While results are promising, limitations include dataset bias, lack of external validation, and scalability challenges in real-world settings. The RL module also depends on timely and accurate patient input, which may be a constraint in low-resource environments. Future iterations should explore electronic health record integration, and so on.

Conclusions

This study presents Nephrosense AI, a novel ML framework designed to transform CKD care through early detection, precise stage prediction, and personalised treatment. Trained on a rich dataset of 4,000 patient records, Nephrosense AI achieved a robust 92% accuracy in CKD stage classification using a random forest classifier, alongside strong regression results for CKD progression prediction, with an R^2 of 0.82 indicating reliable forecasting of kidney function decline over time. The framework's unsupervised clustering component further uncovered meaningful patient subgroups, validated by superior silhouette and Calinski-Harabasz scores, enhancing patient stratification and enabling more targeted intervention strategies. A key innovation lies in the RL module, which adaptively refines treatment recommendations by learning from real-time patient responses. Evaluation showed that RL-driven adjustments led to measurable improvements in treatment efficiency scores, with patients in the adaptive group exhibiting a 15% higher average efficiency compared to those on static treatment protocols, demonstrating the system's capability for personalised, dynamic care optimisation.

What sets Nephrosense AI apart is its hybrid architecture - combining supervised learning, unsupervised clustering, and reinforcement learning - that integrates lifestyle and temporal progression data alongside traditional clinical markers, offering deeper insights into disease mechanisms. This comprehensive approach empowers clinicians and patients with actionable, personalised insights, marking a meaningful step forward in compassionate, data-driven CKD management.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Tejaswi Koul, Yogeshchandra Paranjpe, Aryan Nandanwar, Aniket Mahajan, Asma Parveen I. Siddavatam

Acquisition, analysis, or interpretation of data: Tejaswi Koul, Yogeshchandra Paranjpe, Aryan Nandanwar, Aniket Mahajan

Drafting of the manuscript: Tejaswi Koul, Yogeshchandra Paranjpe, Aryan Nandanwar, Aniket Mahajan

Critical review of the manuscript for important intellectual content: Asma Parveen I. Siddavatam

Supervision: Asma Parveen I. Siddavatam

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. Duff R, Awofala O, Arshad MT, Lambourg E, Gallacher P, Dhaun N, Bell S: Global health inequalities of chronic kidney disease: a meta-analysis. Nephrology Dialysis Transplantation. 2024, 39:1692-1709. 10.1093/ndt/gfae048
2. Guo JJ, Liu Z, Wang P, et al.: Global, regional, and national burden inequality of chronic kidney disease, 1990-

2021: a systematic analysis for the global burden of disease study 2021. Frontiers in Medicine. 2025, 11:1501175. 10.3389/fmed.2024.1501175

3. Chadban S, Arıcı M, Power AJ, et al.: Projecting the economic burden of chronic kidney disease at the patient level (Inside CKD): a microsimulation modelling study. eClinicalMedicine. 2024, 72:102615. 10.1016/j.eclinm.2024.102615
4. Anifah L, Haryanto: Chronic kidney disease severity identification using template matching feature selection statistics based. 2022 International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT), Jember, Indonesia. 2022, 20-24. 10.1109/ICEECIT55908.2022.10030466
5. Anurag, Vyas N, Sharma V, Balla D: Chronic kidney disease prediction using robust approach in machine learning. 2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT), Dehradun, India. 2023, 1-5. 10.1109/CISCT57197.2023.10351277
6. Aruna O, Sameerunnisa S: Chronic kidney disease prediction using data pre-processing techniques. 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India. 2023, 1665-1670. 10.1109/ICSSIT55814.2023.10061025
7. Yildirim P: Chronic kidney disease prediction on imbalanced data by multilayer perceptron. 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), Turin, Italy. 2017, 193-198. 10.1109/COMPSAC.2017.84
8. Tope-Oke A, Badeji-Ajisafe B, Oguntimilehin A, Inyang MV, Aweh O, Abiola O: K-nearest neighbour-based chronic kidney disease prediction system: a case of toxic metals in urine. 2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG), Omu-Aran, Nigeria. 2024, 1-6. 10.1109/SEB4SDG60871.2024.10630163
9. Tummala A, Parvataneni K: Early prediction of chronic kidney disease using AI. 2023 15th Biomedical Engineering International Conference (BMEiCON), Tokyo, Japan. 2023, 1-4. 10.1109/BMEiCON60347.2023.10322036
10. Maurya A, Wable R, Shinde R, John S, Jadhav R, Dakshayani R: Chronic kidney disease prediction and recommendation of suitable diet plan by using machine learning. 2019 International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India. 2019, 1-4. 10.1109/ICNTE44896.2019.8946029
11. Vimala C, Subramani C, Bhagat V, Sravani MV: Analysis of different algorithms to predict chronic kidney disease. 2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC), Bengaluru, India. 202, 1051-1054. 10.1109/IIHC55949.2022.10060311
12. Pu G, Jiang S, Yang Z, Hu Y, Liu Z: Deep reinforcement learning for treatment planning in high-dose-rate cervical brachytherapy. Physica Medica: European Journal of Medical Physics. 2022, 94:1-7. 10.1016/j.ejmp.2021.12.009
13. Saranya G, Sivaraman K: Applying reinforcement learning to optimize treatment plans for chronic disease management. 2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India. 2024, 1-6. 10.1109/ACCAI61061.2024.10601732
14. Schrauben SJ, Apple BJ, Chang AR: Modifiable lifestyle behaviors and CKD progression: a narrative review. Kidney360. 2022, 3:752-778. 10.34067/kid.0003122021
15. Rashed-Al-Mahfuz M, Haque A, Azad A, Alyami SA, Quinn JMW, Moni MA: Clinically applicable machine learning approaches to identify attributes of chronic kidney disease (CKD) for use in low-cost diagnostic screening. IEEE Journal of Translational Engineering in Health and Medicine. 2021, 9:1-11. 10.1109/jtehm.2021.3073629
16. Aldhyani THH, Alshebami AS, Alzahrani MY: Soft clustering for enhancing the diagnosis of chronic diseases over machine learning algorithms. Journal of Healthcare Engineering. 2020, 2020:4984967. 10.1155/2020/4984967
17. CKD (Chronic Kidney Disease) Dataset with Stages. Accessed: April 23, 2025: https://www.kaggle.com/datasets/aryanhirlekar/ckdchronic-kidney-disease-dataset-with-stages/data.
18. Estimated Glomerular Filtration Rate (eGFR). (2022). Accessed: May 27, 2025: https://www.kidney.org/kidney-topics/estimated-glomerular-filtration-rate-egfr.
19. Creatinine Test. (2025). Accessed: May 27, 2025: https://www.testing.com/tests/creatinine/.
20. Blood Urea Nitrogen (BUN) Test. (2022). Accessed: May 27, 2025: https://my.clevelandclinic.org/health/diagnostics/17684-blood-urea-nitrogen-bun-test.
21. Urinalysis: What It Is, Purpose, Procedure, Results & Types. (2024). Accessed: May 27, 2025: https://my.clevelandclinic.org/health/diagnostics/17893-urinalysis.