Literature Review:

CKD is a fatal disease if left undetected as it leads to renal failure, in the worst cases. However, the early diagnosis of CDK can significantly reduce the mortality rate. There are numerous studies available that utilize CKD data and built prediction models depending on the type of data analyzed. But In**[1]**, The processed dataset was trained using different prediction models such as KNN, SVM, RF, and bagging. But the model we have applied for prediction is KNN. The models' performance was estimated to show higher reliability and significance in terms of accuracy, sensitivity, F-measure, specificity, and AUC score. KNN outperformed with an accuracy of 99.50%, sensitivity of 99.2%, precision of 100%, specificity of 98.7%, and F-measure and AUC score of 99.6%. Although the dataset contains all possible attributes that are enough to detect CKD.

in[**2**], new methodology is used which was compared with other machine learning techniques, namely, the K Nearest Neighbor (KNN), SVM, logistic regression, naive Bayes, random forest, and AdaBoost. This study discussed the diagnosis of CKD in accordance with the response of human body. The dataset used for CKD detection in the experiments was obtained from the Kaggle platform and was publicly available. A hybrid model based on Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) was proposed using a two-class data set, which automatically classified CKD. This dataset consisted of thirteen features and one output. If the features showed, CKD was diagnosed. Compared with many well-known machine learning methods, the proposed CNN-LSTM based model obtained a classification accuracy of 99.17%.

in[**3**], the author use the SVM and RF technique to achieved the lowest false-negative rates and test accuracy. A statistical technique, namely, the chi-squared test, is used for the extraction of the least-required set of adequate and highly correlated features to the output. For the model training, a stack of supervised-learning techniques is used for the development of a robust machine-learning model. Out of all the applied learning techniques, support vector machine (SVM) and random forest (RF) achieved the lowest false-negative rates and test accuracy, equal to 99.33%99.33% and 98.67%98.67%, respectively. However, SVM achieved better results than RF did when validated with 10-fold cross-validation.

in[**4**], the author carried out feature optimization, wherein three different feature selection algorithms were applied to find the algorithm most beneficial to extract the important feature for the prediction of Chronic Kidney Disease. As many datasets have imbalanced class, class balancing is needed for increasing the performance of classifier model. In this research SMOTE was used as a class balancer. The highest accuracy of 99.6% was achieved. This dataset is being provided by the UC Irvine Machine Learning Repository and it is available on the UCI website. This dataset contains 400 instances and 24 attributes with 1 target attribute. The target attribute has been labelled in two-class to represent CKD or non-CKD. The dataset was collected from various hospitals. Seven classifier algorithms have been applied in this research such as artificial neural network, C5.0, Chi-square Automatic interaction detector, logistic regression, linear support vector machine with penalty L1 & with penalty L2 and random tree. Although this technique has become reliable for medical treatment.

Literature Summary:

Serial No	*Authors*	*Model*	*Dataset*	*Accuracy*	*Limitations*	*Problem*
1.	Ullah Z. et al	KNN	UCI Machine Learning Repository Among the 24 features,	99.50%	While performance metrics suggest reliability, real-world applicability needs confirmation through external validation on diverse datasets. Independent sources' validation is essential to ensure the models' robustness.	Enhancing Early Detection of Chronic Kidney Disease: A Prediction-Based Approach for Rapid and Cost-Effective Decision-Making
2.	Yildiz et al	CNN-LSTM	UCI Machine Learning Repository	99.17%.	In future studies, new models can be developed to detect CKD on a multicenter data set with more data.	Extract Accurate Diagnosis Using Feature-Rich Data"
3.	Swain, D. et al	SVM and RF	UCI Machine Learning Repository	99.33% and 98.67 %	achieving higher accuracy with fewer features acknowledges and improvement	Approaches for Accurate Diagnosis using Publicly Available Data
4.	Chittora, P, et al	Seven classifier	UC Irvine Machine	99.6%.	Evaluating machine	integration of models

		algorithms	Learning Repository		learning algorithm and deep neural network for enhanced accuracy	into clinical practice, with extensive validation studies across diverse patient

References:

1. Ullah, Z. and M. Jamjoom, *Early detection and diagnosis of chronic kidney disease based on selected predominant features.* Journal of Healthcare Engineering, 2022. **2023**.
2. Yildiz, E., et al., *Diagnosis of chronic kidney disease based on CNN and LSTM.* Acadlore Trans. Mach. Learn, 2023. **2**(2): p. 66-74.
3. Swain, D., et al., *A Robust Chronic Kidney Disease Classifier Using Machine Learning.* Electronics, 2023. **12**(1): p. 212.
4. Chittora, P., et al., *Prediction of chronic kidney disease-a machine learning perspective.* IEEE Access, 2021. **9**: p. 17312-17334.