International Journal of Research in Science and Technology

Volume 12, Issue 01: January - March 2025

ISSN 2394 - 9554

Prediction of Chronic Kidney Disease Risk using AI Techniques Considering Various Environmental and Genetic Factors

Akshitha D. Gopikrishnan[1] and Dr. S. Prabhakar Karthikeyan[2]

1.Research Scholar, Rock Hill High School, Frisco, Texas, USA

2.Professor, Vellore Institute of Technology, Vellore, Tamilnadu India

*** Akshitha.Gopikrishnan@gmail.com**

Abstract

Chronic Kidney Disease (CKD) is a global crisis affecting 10-15% of the world population and affects a few million people worldwide. As CKD expected to rise as a major health issue, it heavily burdens the strained healthcare systems, particularly in low- and middle-income countries. The economic impact associated with the treatments and long-term care, will further challenge global health systems. Despite the ever-increasing prevalence, a considerable proportion of affected individuals remains undiagnosed; this leads to premature mortality. Undetected CKD progresses through stages, ultimately leading to kidney failure, which requires dialysis or a transplant as an ultimate care. When this get linked to cardiovascular diseases it significantly impacts patient's quality of life, leading into long-term disability. The early detection and personalized management of CKD plays an important role in preventing disease progression and improving patient prognosis. This research aims to address a critical gap by developing a robust algorithm that calculates an individual's risk level of CKD incorporating both genetic predispositions and environmental factors. By leveraging machine learning techniques and with extensive data analysis, the proposed tool will enable healthcare professionals worldwide to deduct the high-risk patients at an early stage. This research resulted in the development of a detailed, yet user-friendly, risk assessment tool that facilitates early detection and prevention of kidney disease. Enabling healthcare professionals by providing with a tool that can be reliable and accessible method for identifying the risk based on several influencing factors ultimately improves disease management and patient outcomes. Implementation of such tools within health systems will translate into a significant impact on the reduction of the burden of CKD worldwide.

(One blank line)

Keywords: *CKD, Machine Learning Algorithms, Risk Prediction Models, XGBoost Classifier, Feature Importance Analysis, SHAP Values, Genetic and Environmental Factors, Early Detection and Prevention, Personalized Risk Assessment, Healthcare Decision Support Systems*

1. INTRODUCTION

CKD is a worldwide health burden that seriously affects millions of people, causing substantial morbidity, mortality, and costs of care. Effective disease management widely depends on early detection and prevention of chronic conditions. However, there is a lack of accuracy and gap in the availability of predictive tools that combine both genetic and environmental factors for assessing the risk. This research develops a machine-learning algorithm that calculates the risk percentage for developing kidney disease based on the comprehensive analysis of genetic predispositions and environmental factors.

Major Contributions of this study

- This research identifies a previously unexplored factors that correlates with disease severity in CKD patients and offers early disease detection.
- Developed comprehensive dataset and the machine learning model using XGBoost fine-tuned with GridSearchCV achieved excellent performance in predicting the risk of kidney disease.
- Through this study, it has been demonstrated that XGBoost could significantly enhance early detection of CKD and offers a cost-effective, scalable solution for improving patient care and preventing the disease progression.
- This study conducted compared the analysis of **XGBoost** with other popular machine learning algorithms. The XGBoost model outperformed both in terms of key evaluation metrics as discussed in this paper.

1.1 Global Burden of Chronic Kidney Disease

It's estimated around 700 million people worldwide have CKD, representing about 10% of the global population. There are several genetic and environmental factors contributing to CKD, among which common factors include diabetes, hypertension, and obesity. For instance, diabetes alone is responsible for about 30-40% of CKD cases worldwide (Friedman, 2019). People with CKD have a 50% higher risk of death from heart disease compared to those without kidney disease (Ahmed et al., 2022). Globally, around 2-3 million people require kidney transplantation or dialysis annually (Unger et al., 2020). Awareness of the disease is often low, leading to late-stage diagnosis and treatment (Gregorich et al., 2023).

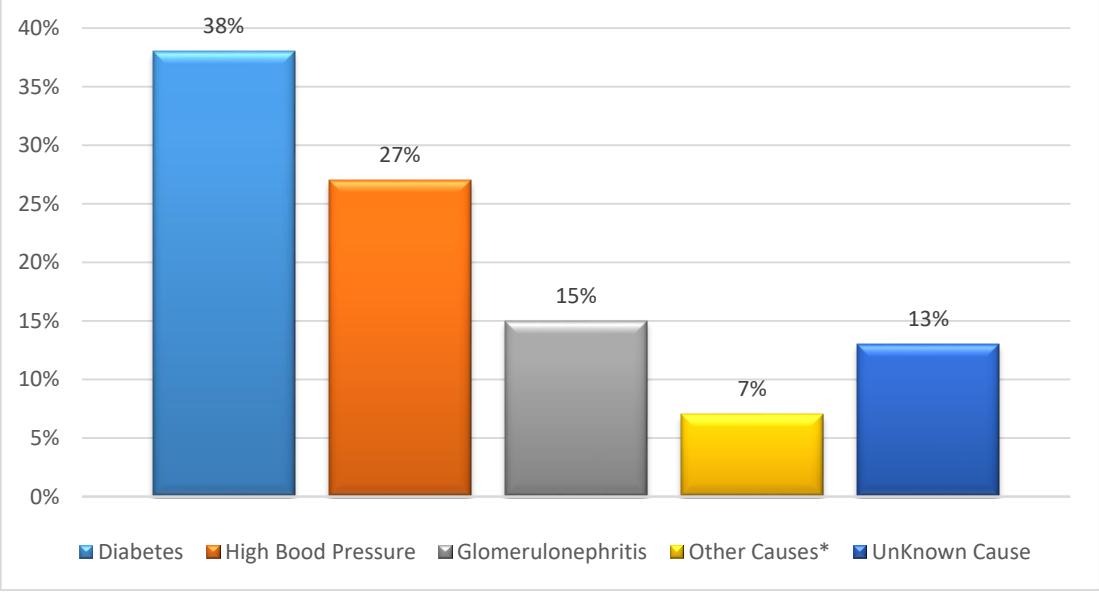

Figure 1.1: Source: US Renal Data System. Reported causes of end-stage kidney disease in the US (N=807,920, all ages, 2020). Includes polycystic kidney disease, among other causes.

1.2Complexity of Genetic and Environmental Interplay

As CKD is a global health burden and its prevalence is increasing day by day, with the special increase in the prevalence rate of hypertension, diabetes, and obesity, the need for overall risk assessment tools is critical. The early detection coupled with appropriate personalized treatment forms the cornerstone in reducing the burden of kidney diseases (Ahmed et al., 2022). In fact, Studies have shown that earlier intervention in the course of the disease leads to a better prognosis, as it results in fewer patients progressing to End-stage Renal Disease (ESRD). Such innovative thinking could include the integration of machine learning and AI to predict the progression of CKD in a way that allows health systems to prioritize interventions (US Centers for Medicare and Medicaid Services, 2024). Integrating machine learning with AI into CKD risk prediction can help prioritize timely interventions within healthcare systems.

Public health campaigns that encourage healthy eating, staying active, and regular kidney checkups can help reduce the number of people who don't know they have kidney problems. With more focus and funding on preventing kidney disease, along with better tools to predict who is at risk, we can lower the number of cases and their impact around the world. (Francis et al., 2024; Zhang, Fang, & Tran, 2023).

Studies suggest that individuals with a family history of kidney disease may have up to a 50% higher risk of developing chronic kidney disease compared to those without such a history (Saran et al., 2022). Genetic predisposition to kidney disease is the major contributing factor, accounting for about 75% of the overall risk. The most prominent genetic factors are family history of kidney disease, gender, age from 40-80, race, and other diseases such as diabetes and high blood pressure (Friedman, 2019; Harasemiw et al., 2019). Newer research identifies specific genetic polymorphisms associated with kidney disease, thereby making genetics play an even greater role in predicting disease makeup and their management (Francis et al., 2024).

Environmental and social determinants independently contribute to a significant proportion, about 25%, of the risk for kidney disease. Major environmental risk factors include unhealthy dietary behaviors—for example, high intake of sodium or poor hygiene practices, exposure to toxic environments—metals or air pollutants—and access to quality healthcare (Research on Environment and Kidney Disease, 2023). This involves dietary patterns in food intake - for instance, in processed foods—and high levels of potassium. Other variables that were measured in this study include the degree of physical activity, quality of water, exposure to occupational hazards, socio-economic status, degree of social support from the community, utilization of healthcare facilities, smoking and drinking habits, medication, BMI, and stress levels (Saran et al., 2022; Social Determinants of Health and Kidney Disease, 2023). The genetic and environmental interplay at the initiation and progression of kidney disease is complex, and its early detection, prevention, and efficient management require an understanding of these factors.

The study below shows that CKD was more common in persons aged 65 years or older (34%) than either of the other two age groups, aged 45–64 years (12%) and 18–44 years (6%). In addition, CKD was slightly more common among women (14%) than men (12%). CKD was more common among non-Hispanic Black adults (20%) than either non-Hispanic Asian adults (14%) or non-Hispanic White adults (12%). CKD was found in approximately 14% of Hispanic adults.

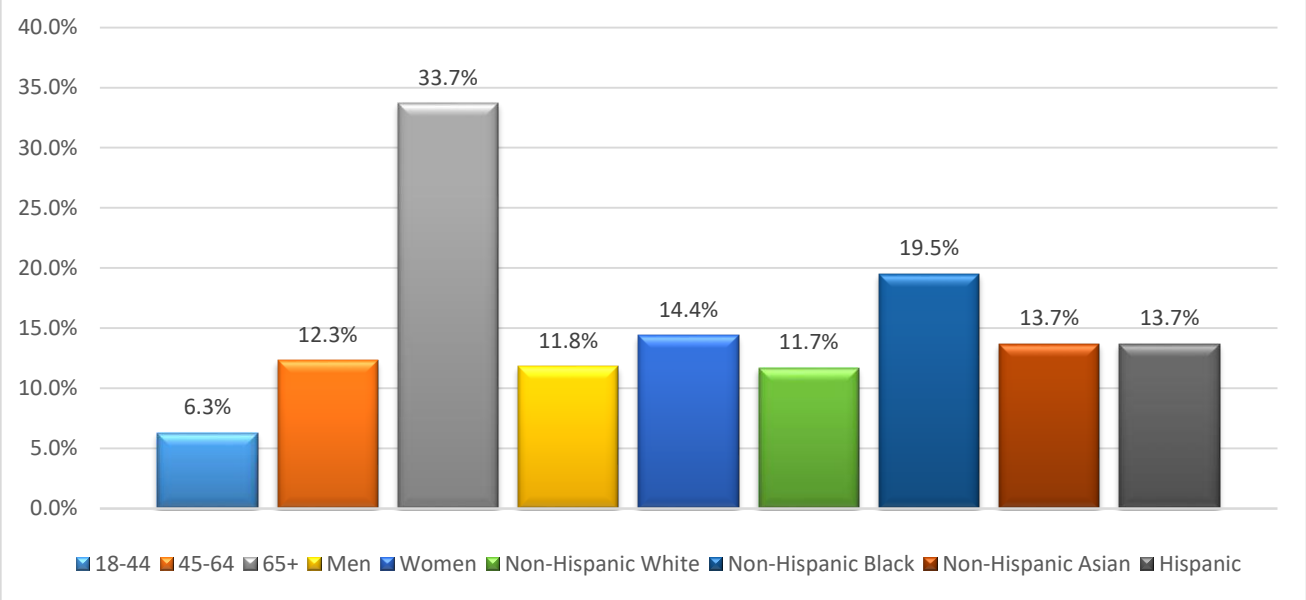

Figure 1.4: Source: Centers for Disease Control and Prevention. Chronic Kidney Disease in the United States, 2023. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2023. Percentage of US Adults Aged 18 Years and Older With CKD, by Age, Sex, and Race/Ethnicity in the US.

2. Building an Accurate and Comprehensive Risk Prediction Model

To build a robust predictive model, extensive study was conducted using a comprehensive CKD dataset from Kaggle, that includes various factors such as patient demographics, medical history, and biomarkers. The GridSearchCV-optimized XGBoost classifier demonstrated strong performance, achieving high accuracy and robustness in predicting CKD risk, with metrics like AUC-ROC reaching 0.98, indicating its effectiveness on the dataset used. Feature importance was interpreted using SHapley Additive exPlanations (SHAP) values, using the key predictors such as albumin levels, hemoglobin levels, blood glucose, and age.

Even though individual genetic and environmental factors have been studied over a very long period, there has been a failure in the availability of one integrated tool with these various factors to provide an individualized risk assessment. This research experiment aims to fill that gap by developing an algorithm that will allow for risk percentage calculations of kidney disease based on a combination of both genetic and environmental factors. The results can further be analyzed to assess the progression rate of kidney disease over time based on the calculated risk.

2.2Personalized Care for Early Interventions

The primary goal of this experiment is to create a robust machine-learning model that can accurately estimate the risk factors associated with renal disorders by analyzing a range of genetic and environmental influences. The intention behind this model is to help users figure out their possible risks so that they can start personalized early treatment or at least take precautions against.

This research leverages advanced algorithms with comprehensive data sets to enable the model to identify patterns and correlations that may not be immediately apparent through traditional assessment

methods (Zhang, Fang, & Tran, 2023). Furthermore, the model design can be further enhanced with user-friendly interfaces to ensure accessibility, enabling individuals to input their personal data easily through healthcare providers, Health Systems, and Insurance Companies. Ultimately, these type of initiatives aims to enhance the overall management of renal health, reduce the incidence of advanced kidney disease, and improve outcomes for at-risk populations (Francis et al., 2024; Harasemiw et al., 2019).

The development of this algorithm and risk assessment tool aims to facilitate early detection and personalized preventive measures for kidney disease. With improved patient outcomes for those diagnosed and a contribution made toward general knowledge of the risk factors that surround kidney disease, this project delivers a comprehensive, yet easy-to-use, risk assessment tool (Saran et al., 2022).

As many as 9 out of every 10 adults with CKD do not know they have the disease, underscoring an urgent need for effective risk assessment and awareness tools (Gregorich et al., 2023).

3. Methods and Tools for CKD Risk Prediction

3.1 Data Collection

The dataset used in this study is the CKD dataset from Kaggle. This comprehensive dataset includes a wide range of features such as patient demographics, medical history, and various biomarkers crucial for understanding CKD risk. Notable features include age, blood pressure, sugar levels, sodium levels, potassium levels, diabetes status, family history, exercise frequency, and overall health metrics. This rich dataset serves a robust foundation for developing a predictive model that can discern patterns indicative of kidney disease risk. By harnessing these diverse data points, this model enhances early detection and prevention strategies for CKD.

3.2 Tools and Libraries Used

To conduct this experiment effectively, a wide range of tools and libraries was used to ensure efficient data analysis and machine learning processes. Python is the primary programming language chosen due to its versatility and robust support for data analysis and machine learning. For data manipulation and pre-processing, Pandas was utilized, while numerical operations and array handling were carried out using NumPy. To implement the gradient boosting classifier **XGBoost** was selected due to its high-performance rate in classification tasks. **Scikit-learn** provides essential tools for machine learning such as data splitting and hyperparameters tuning functions. For data visualization, **Matplotlib** and **Seaborn** were used to create insightful graphs and plots, help to highlight critical key patterns and relationships within the analyzed dataset. Additionally, the **SHAP (SHapley Additive exPlanations)** method was integrated to interpret the model's predictions. By quantifying each feature's contribution to the output, SHAP ensured transparency and interpretability, making the model's decisions more comprehensible and actionable.

3.3 Data Preprocessing and Feature Engineering

The initial phase of the research experiment involved data collection and preprocessing. The CKD dataset was sourced from Kaggle using Kaggle API. The downloaded zip file was programmatically extracted and imported into a Pandas DataFrame. This step was undertaken to ensure the dataset was structured and readily accessible for downstream data wrangling and for additional processing.

Next step focused on data cleansing to ensure the dataset's quality. Subsequently, the process of data cleansing was conducted with the elimination of missing values in the dataset. This was essential to maintain data integrity and avoid any potential biases or inaccuracies during model training. Categorical

variables, such as ‘rbc’ (red blood cells), ‘pc’ (pus cells), ‘pcc’ (pus cell clusters), ‘ba’ (bacteria), ‘htn’ (hypertension), ‘dm’ (diabetes mellitus), ‘cad’ (coronary artery disease), ‘pe’ (pedal edema), ‘ane’ (anemia), and ‘appet’ (appetite), were encoded into binary values through one-hot encoding or label encoding techniques to enable effective integration with machine learning algorithms.

	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd
3	3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	...	35	7300	4.6	no	no	no	good	no	no	ckd
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
395	395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	47	6700	4.9	no	no	no	good	no	no	notckd
396	396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	54	7800	6.2	no	no	no	good	no	no	notckd
397	397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	49	6600	5.4	no	no	no	good	no	no	notckd
398	398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	51	7200	5.9	no	no	no	good	no	no	notckd
399	399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	53	6800	6.1	no	no	no	good	no	no	notckd

400 rows x 26 columns

Figure 3.1: Dataset shows sample vital information and various influencing factors.

To suppress the effort of potential outliers on model performance and accuracy, the interquartile range (IQR) method was deployed. Any data points outside 1.5 times the IQR above the third quartile or below the first quartile were considered as outliers and excluded from the dataset ensuring that the model trained on clean, reliable data. Additionally, normalization and scaling were performed on the numerical features using Min- Max scaling, bringing all features to a consistent range of [0, 1], which is important for improving the performance of algorithms like XGBoost. This normalization process is essential when working with gradient-based models, such as XGBoost, as it improves convergence speed and ensures that the model treats each feature with equal importance. Following preprocessing, exploratory data analysis (EDA) was conducted using statistical and visual methods to gain deep insights. A correlation matrix was made to establish how each variable is related to all others, and descriptive statistics were obtained that indicated the average, spread, and distribution shape for this data set. By conducting this procedure, it became easier to pinpoint possible kidney disease predictors as well as decide on the features that could be used after having identified them first.

ISSN 2394 - 9554

```
from xgboost import XGBClassifier
from sklearn.model_selection import GridSearchCV
parameters = {
    'n_estimators': [0, 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50],
    'max_depth': [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20]
}
model = XGBClassifier(objective='binary:logistic', learning_rate=0.1)
grid = GridSearchCV(estimator=model, scoring='accuracy', cv=2, param_grid=parameters)
grid.fit(x_train, y_train)
```

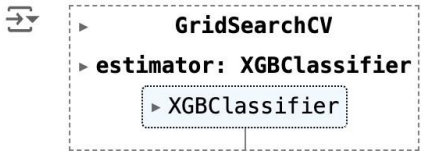

Figure 3.2: Importing **XGBoost** & GridSearchCV view

During feature selection, the variance threshold method from Scikit-learn was applied to eliminate features with low variance, ensuring that the model would focus only on the most relevant attributes.

This dimensional reduction technique ensured that only the most informative features remained in the dataset, thereby improving computational efficiency and reducing the complexity of the model without sacrificing predictive power and accuracy.

4.Optimizing Model Performance: Feature Selection and Data Splitting

To optimize the predictive accuracy for kidney disease, a feature identification step was deployed where features with less variance were eliminated through the use of variance threshold. This powerful technique ensured that only the most critical influential features are remained. This allowed the model to focus on the key attributes that contribute significantly to the prediction task with greater accuracy. By removing less impactful variables, the model's efficiency was greatly improved, and its ability to generalize was maximized.

The dataset was split into training and testing subsets using an 80-20 ratio, with 80% allocated for model training and 20% reserved for testing. This approach ensured that a major portion of the data was used for training, while a distinct, unseen block was set aside for performance evaluation. The separation of training and testing data allowed for a fair estimate of the model's generalizability and ensured that the model's performance was evaluated on data it had not encountered during training, reducing the risk of overfitting.

```
[ ]  df_y = df["classification"]
     df_X = df.drop(columns = ["classification", "id"])
```

```
[ ]  from sklearn.feature_selection import VarianceThreshold
     df_X = VarianceThreshold(threshold=0.1).fit_transform(df_X)
```

Figure 2.3: Snapshot of variance threshold being used

5.Results and Discussion

The developed machine learning model in this study showed a high degree of accuracy and delivered impressive results in predicting the risk of kidney disease. After going through intensive preprocessing and feature selection stages, the **XGBoost classifier**, fine-tuned with **GridSearchCV**, achieved excellent performance on the test data set, with high accuracy. While no model is perfect, the results indicate that the model was able to make highly accurate predictions on unseen data, demonstrating its ability to generalize well and effectively identify kidney disease risk level.

To evaluate the effectiveness of the XGBoost model, we compared its performance to other common algorithms such as Random Forest and Support Vector Machines (SVM). The XGBoost model, with its hyperparameter optimization, consistently outperformed both Random Forest and SVM in terms of accuracy, precision, recall, and F1 score. While Random Forest showed competitive results, especially in handling imbalanced data, it did not achieve the same level of optimization as XGBoost. The SVM, despite being effective in certain binary classification tasks, lagged behind XGBoost in performance for this dataset due to its sensitivity to the scaling of input features and hyperparameter tuning. All in all, **XGBoost** was the best choice for predicting kidney disease risk. To make sure we were using the most useful features, **Variance Threshold** technique was used to remove unnecessary data. We also used **SHAP (SHapley Additive exPlanations)** values to figure out which features had the biggest impact on our model's predictions. Some of the key factors that showed up as important were **albumin**, **hemoglobin**, **blood glucose**, and **age**, which match what we already know about kidney disease. According to the following chart; Albumin, hemoglobin, blood glucose, and age were some highly contributing factors to the development of kidney diseases. (Francis et al., 2024)

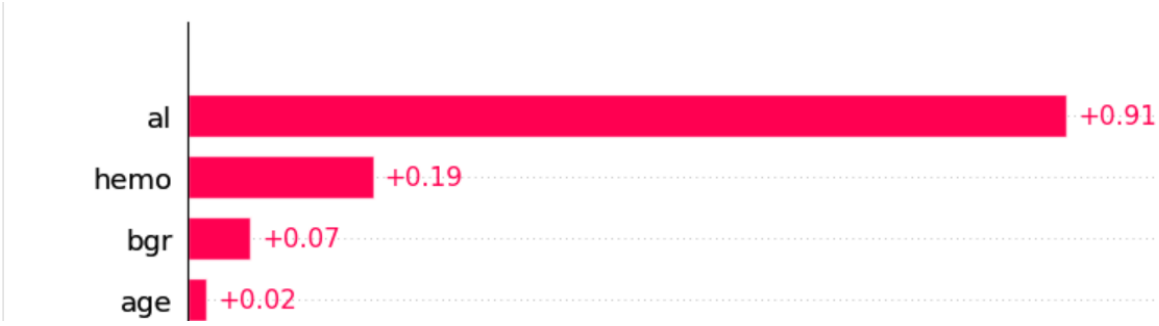

Figure 3.1: Snapshot of results of the SHAP Plot Demonstrating the Factors that Contributed the Most to the Development of Chronic Kidney Disease With 0 Being Not Affected and 1 Being Affected.

The SHAP analysis revealed that features such as blood glucose and age played a critical role in the model's predictions. Specifically, the higher the blood glucose level, the higher the likelihood of chronic kidney disease, highlighting the importance of monitoring diabetes as a risk factor. Age was another key predictor, with individuals over the age of 60 being at a significantly higher risk of developing kidney disease. These insights from SHAP values provide valuable understanding of the risk factors, further reinforcing the relevance of monitoring these health indicators in the prevention and early detection of kidney disease (Ahmed et al., 2022). Because of these factors, we can interpret which pre-existing conditions lead to a higher risk of kidney disease. For example, high levels of blood glucose are an indicator of diabetes, which causes 47% of new kidney disease cases. Additionally, Chronic Kidney

Disease typically occurs in people ages 60 or older. The level of albumin is important, as lower levels of the blood plasma are a possible result of kidney disease. The final factor that significantly contributed to the risk of chronic kidney disease is the levels of hemoglobin. Lower levels of hemoglobin can lead to anemia, which is a possible complication of kidney disease (Francis et al., 2024).

Overall **XGBoost** proved to be a powerful tool for predicting kidney disease risk. This study shows that **XGBoost** could be a game-changer for predicting kidney disease. While it's not perfect, it outperforms traditional methods and could help doctors detect kidney disease early, allowing for better treatment before it becomes severe. It offers a practical, cost-effective way to improve patient care and prevent the serious consequences of late-stage kidney disease.

6. Conclusion

To conclude, this research suggests, adding more types of data, like genetic information and environmental factors would help make the model more accurate and robust. To make the kidney disease prediction model even better, incorporating more diverse datasets like gathering data from different populations groups, different regions, socioeconomic backgrounds, and medical histories would be able to give more personalized predictions. Incorporating diverse datasets, such as data from varying geographical regions, socioeconomic statuses, and medical histories, would enhance the model's ability to account for a wide range of variables, providing more precise and personalized risk predictions. With the data inclusive of all these aspects, a holistic understanding of the risk factors and their interaction could be taken into account for more precise predictions (Friedman, 2019; Zhang, Fang, & Tran, 2023). These additional factors would allow the model to provide a deeper understanding of the risk factors, which would improve its ability to predict kidney disease risk at a greater accuracy for personalized treatments.

Another way to improve the model is by fine-tuning it even more and testing out different machine learning techniques and with vast amount of diverse data. Combining multiple models in an ensemble approach could help increase accuracy because it leverages the strengths of different algorithms, making the predictions more reliable. The goal is to turn the model into a user-friendly tool that people can use to input their health data and get personalized advice. This tool could help profile the patient and offer dietary tips like reducing sodium intake or suggest exercise routines based on individual risk factors. This would put people to make healthier choices and possibly prevent the need for expensive treatments in the future. Ultimately, this could help reduce healthcare costs by allowing for earlier interventions, all from the comfort of a person's own home. For this to be widely used, there needs to be an easy-to-use interface for both healthcare professionals and patients. This would allow doctors to input data, track patient health, and provide actionable recommendations. A dashboard for healthcare providers would be helpful, too, allowing them to interpret the data in real-time, making it a valuable tool in clinical settings.

Most importantly, the model should be able to learn and update itself with new data regularly. This ensures that it stays relevant and accurate as new trends, research, and methods emerge. A continuous learning system would allow the model to adapt to changes in health trends, and incorporating feedback from both patients and healthcare providers would help it improve over time. This feedback loop would help keep the model effective, ensuring that it keeps improving and providing real-world value.

Acknowledgments*:*

I would like to thank Dr. S. Prabhakar Karthikeyan, Professor, PDF (CPRI), Senior Member (IEEE), for his invaluable guidance and support throughout the course of this research. His expertise in research

publications greatly enhanced the depth and quality of this study.

7.Reference

1. Ahmed, S., Mothi, S. S., Sequist, T., Tangri, N., Khinkar, R. M., & Mendu, M. L. (2022). The Kidney Failure Risk Equation Score and CKD Care Delivery Measures: A Cross-sectional Study. *Kidney Medicine, 4*(1). https://doi.org/10.1016/j.xkme.2021.08.010
2. ESRD Treatment Choices (ETC) Model | CMS. (n.d.). *Centers for Medicare & Medicaid Services*. https://www.cms.gov/priorities/innovation/innovation-models/esrd-treatment-choices-model
3. Kidney Care Choices (KCC) Model | CMS. (n.d.). *Centers for Medicare & Medicaid Services*. https://www.cms.gov/priorities/innovation/innovation-models/kidney-care-choices-kcc-model
4. Francis, A., Harhay, M. N., Ong, A. C. M., Tummalapalli, S. L., Ortiz, A., Fogo, A. B., Fliser, D., Roy-Chaudhury, P., Fontana, M., Nangaku, M., Wanner, C., Malik, C., Hradsky, A., Adu, D., Bavanandan, S., Cusumano, A., Sola, L., Ulasi, I., & Jha, V. (2024). Chronic kidney disease and the global public health agenda: an international consensus. *Nature Reviews Nephrology*, 1–13. https://doi.org/10.1038/s41581-024-00820-6
5. Friedman, D. J. (2019). Genes and environment in chronic kidney disease hotspots. *Current Opinion in Nephrology and Hypertension, 28*(1), 87–96. https://doi.org/10.1097/MNH.0000000000000470
6. Gregorich, M., Kammer, M., Heinzel, A., Böger, C. A., Eckardt, K., Lambers Heerspink, H., Jung, B., Mayer, G., Meiselbach, H., Schmid, M., Schultheiß, U. T., Heinze, G., & Oberbauer, R. (2023). Development and Validation of a Prediction Model for Future Estimated Glomerular Filtration Rate in People With Type 2 Diabetes and Chronic Kidney Disease. *JAMA Network Open, 6*(4), e231870. https://doi.org/10.1001/jamanetworkopen.2023.1870
7. Harasemiw, O., Drummond, N., Singer, A., Bello, A., Komenda, P., Rigatto, C., Lerner, J., Sparkes, D., Ferguson, T. W., & Tangri, N. (2019). Integrating Risk-Based Care for Patients With Chronic Kidney Disease in the Community: Study Protocol for a Cluster Randomized Trial. *Canadian Journal of Kidney Health and Disease, 6*, 205435811984161. https://doi.org/10.1177/2054358119841611
8. New Model Predicts Cardiovascular Risk Among CKD Patients. (n.d.). *Penn Medicine*. Retrieved August 1, 2023, from https://www.pennmedicine.org/news/news-releases/2023/july/new-model-predicts-cardiovascular-risk-among-ckd-patients
9. Population Health and Chronic Kidney Disease: Improving Outcomes with Technology (2023). *Harvard Public Health Review, 45*, 67–80. https://www.hsph.harvard.edu/public-health-review/2023/08/population-health-and-chronic-kidney-disease-improving-outcomes-with-technology/
10. Research on Environment and Kidney Disease. (2023). *National Institute of Environmental Health Sciences*. https://www.niehs.nih.gov/health/topics/conditions/kidney-disease/index.cfm
11. Saran, R., Robinson, B., Abbott, K. C., Bragg-Gresham, J., Balkrishnan, R., Dietrich, X., & Pisoni, R. (2022). US Renal Data System 2022 Annual Data Report: Epidemiology of Kidney Disease in the United States. *American Journal of Kidney Diseases, 79*(4), A8–A18. https://doi.org/10.1053/j.ajkd.2022.01.012
12. Social Determinants of Health and Kidney Disease. (2023). *Kidney.org*. https://www.kidney.org/atoz/content/social-determinants-health-and-chronic-kidney-disease
13. Tangri, N., Grams, M. E., & Inker, L. A. (2020). Risk Stratification in Chronic Kidney Disease: Challenges and Opportunities. *Nephrology Dialysis Transplantation, 35*(2), 219–225. https://doi.org/10.1093/ndt/gfz153

14. Unger, T., Borghi, C., Charchar, F. J., Khan, N. A., Poulter, N. R., Prabhakaran, D., Schlaich, M. P., Stergiou, G. S., Tomaszewski, M., Touyz, R. M., Whelton, P. K., & Williams, B. (2020). 2020 International Society of Hypertension Global Hypertension Practice Guidelines. *Hypertension, 75*(6), 1334–1422. https://doi.org/10.1161/hypertensionaha.120.15026
15. US Centers for Medicare and Medicaid Services. (2024). Innovative Care Models to Improve Outcomes in CKD. *Centers for Medicare & Medicaid Services*. https://www.cms.gov/outreach-and-education/medicare-learning-network/mlnproducts/innovative-care-models-ckd
16. Zhang, Y., Fang, C., & Tran, M. (2023). Social and Environmental Predictors of Chronic Kidney Disease in Urban Settings. *Environmental Research, 224*, 115027. https://doi.org/10.1016/j.envres.2023.115027
17. Jha, V., Garcia-Garcia, G., Iseki, K., Li, Z.-H., Naicker, S., Plattner, B., & Yang, C.-W. (2013). Chronic kidney disease: global dimension and perspectives. *The Lancet, 382*(9888), 260–272. https://doi.org/10.1016/S0140-6736(13)60687-X
18. Luyckx, V. A., & Brenner, B. M. (2005). The impact of the changing epidemiology of chronic kidney disease on public health. *Nature Clinical Practice Nephrology, 1*(7), 3–10. https://doi.org/10.1038/ncpcardio0394
19. Go, A. S., Chertow, G. M., Fan, D., McCulloch, C. E., & Hsu, C.-Y. (2014). Chronic kidney disease and the risk of death, cardiovascular events, and hospitalization. *The New England Journal of Medicine, 351*(13), 1296–1305. https://doi.org/10.1056/NEJMoa041031
20. Xie, Y., Yang, S., & Lou, J. (2021). Global, regional, and national burden of chronic kidney disease, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet, 395*(10225), 709–717. https://doi.org/10.1016/S0140-6736(20)32514-5
21. Levey, A. S., & Coresh, J. (2012). Chronic kidney disease. *The Lancet, 379*(9811), 165–180. https://doi.org/10.1016/S0140-6736(11)60178-5
22. United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World Population Prospects 2019: Highlights*. United Nations. https://population.un.org/wpp/Publications/Files/WPP2019_Highlights.pdf
23. Centers for Disease Control and Prevention (CDC). (n.d.). *CKD National Facts*. https://www.cdc.gov/kidney-disease/php/data-research/?CDC_AAref_Val=https://www.cdc.gov/kidneydisease/publications-resources/CKD-national-facts.html.