Lifestyle-Driven CKD Risk Prediction in Diabetes Using AutoGluon

CHIH-HSUAN HUANG
Department of Biochemistry and Molecular Cell Biology, School of Medicine, College of Medicine
Taipei Medical University
Taipei, Taiwan
chihhsuanhuang086@gmail.com

YU-TING WU
School of Medicine, College of Medicine
Taipei Medical University
Taipei, Taiwan
b101113003@tmu.edu.tw

YEN-HSI CHEN
College of Electrical Engineering and Computer Science, Department of Electrical Engineering
National Tsing Hua University
Hsinchu, Taiwan
nthu113061240@gmail.com

TE-CHAO FANG*
M.D. PhD, Division of Nephrology, Department of Internal Medicine,
Taipei Medical University Hospital
Taipei Medical University
Taipei, Taiwan
Division of Nephrology, Department of Internal Medicine, School of Medicine, College of Medicine
Taipei Medical University
Taipei, Taiwan
Taipei Medical University Research Center of Urology and Kidney (RCUK)
Taipei Medical University
Taipei, Taiwan
fangtechao@gmail.com

FU-DER MAI*
Department of Biochemistry and Molecular Cell Biology, School of Medicine, College of Medicine
Taipei Medical University
Taipei, Taiwan
fordmai@tmu.edu.tw

*Author to whom correspondence should be addressed:
fordmai@tmu.edu.tw, fangtechao@gmail.com

***Abstract*—Chronic kidney disease (CKD) is a major global health burden, characterized by a progressive decline in renal function over time. Compared with reliance on physiological data, we adopt lifestyle factors for risk prediction. In patients with diabetes, chronic hyperglycemia further predisposes to glomerular hyperfiltration and renal injury. In this study, we utilized AutoGluon, an automated machine learning (AutoML) framework, to improve prediction of CKD risk in diabetic patients. The platform also provides a user-friendly graphical user interface (GUI), which simplifies workflow and improves efficiency. The dataset was obtained from BIRDEM General Hospital in Bangladesh and included 22 key features for model training, encompassing demographic variables (e.g., gender, body mass index [BMI]), comorbidities, and lifestyle factors (including diet, physical activity, smoking, and water intake), as well as urinary status. We compared three models—AutoGluon, XGBoost, and TabPFN—evaluating their performance using metrics such as accuracy, area under the ROC curve (ROC-AUC), precision, and recall. The results demonstrated that AutoGluon and TabPFN outperformed XGBoost. Specifically, the F1-scores were 0.8948, 0.9642, and 0.8718, respectively; recall values were 0.8264, 0.9617, and 0.9267; precision values were 0.9755, 0.9666, and 0.8231; accuracy values were 0.9813, 0.9648, and 0.8638; and ROC-AUC values were 0.9778, 0.9648, and 0.8638. Despite the superior numerical performance of TabPFN, its interpretability was limited. Notably, it suggested that longer duration of diabetes reduced the likelihood of CKD, a finding that is contradicted established clinical knowledge. Therefore, AutoGluon was selected as the primary predictive model. These findings highlight AutoGluon's potential to support early intervention and disease prevention in clinical practice.**

Keywords—Chronic Kidney Disease, Lifestyle Features, AutoGluon, Diabetes

I. INTRODUCTION

Chronic kidney disease (CKD) results from a sustained lose in renal function and is commonly manifested by symptoms such as edema, hypertension, and anemia [1]. Populations with diabetes mellitus (DM), gout, obesity, or high-salt dietary habits are at elevated risk for kidney disease.

As diabetes progresses, it induces systemic complications, and prolonged hyperglycemia impairs renal function [2]. Chronic hyperglycemia damages tubuloglomerular feedback, leading to dilation of the afferent arteriole. Simultaneously, activation of the RAAS system enhances the effect of angiotensin II on the efferent arteriole. This dual mechanism results in persistent glomerular hyperfiltration, which, over time, contributes to renal damage and increases the risk of subsequent complications. As renal injury is irreversible, patients in the early stages can only slow disease progression through pharmacological and dietary interventions, whereas those in the advanced stages often require long-term hemodialysis or await kidney transplantation. Implementing preventive measures at an early stage can substantially mitigate disease progression.

In Taiwan, the prevalence of chronic kidney disease in 2022 was approximately 11% (2.6 million individuals) [3]. Among patients with diabetes, about 33% develop kidney disease within 10 years. Both figures are higher compared to neighboring countries, including Korea, Japan, and China. Projections suggest that by 2027, the prevalence will rise to 12.4% (3 million individuals). The annual total expenditure of the National Health Insurance is estimated to increase from NT$51.96 billion in 2022 to NT$62.18 billion in 2027. As kidney transplantation remains uncommon at present, patients with kidney disease largely depend on pharmacological treatment and renal replacement therapies (hemodialysis and

peritoneal dialysis). This situation is expected to impose a substantial burden on the National Health Insurance system, potentially affecting the allocation of healthcare resources for other patients.

The application of artificial intelligence (AI) in the management of chronic kidney disease (CKD) has been expanding rapidly, with applications ranging from early disease detection, risk stratification, prognosis prediction, and automated image analysis (e.g., ultrasound or pathology) to personalized treatment planning and clinical decision support systems. Beyond their applications, previous studies have shown that AI models deliver tangible benefits in CKD care—enhancing patient prognosis, enabling more personalized treatment strategies, and improving the efficiency of healthcare resource allocation [4].

In addition, AI has also been applied to the prediction of complication risks among CKD patients. For instance, one study employed Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis to identify eight key predictive factors—including age, history of hypertension, gender, use of antiplatelet agents, high-density lipoprotein (HDL) levels, serum sodium concentration, 24-hour urinary protein levels, and estimated glomerular filtration rate (eGFR)—to estimate the risk of cardiovascular disease (CVD) in patients with CKD [5]. Such approaches facilitate the early identification of high-risk individuals and enable timely interventions.

In recent years, a growing number of studies have proposed and documented methods for predicting CKD risk among patients with diabetes. One such study focused on patients with type 2 diabetes mellitus [6], utilizing logistic regression (LR) and incorporating seven routinely available clinical features: age, body mass index (BMI), eGFR, serum creatinine, albumin, glucose, and hemoglobin A1c (HbA1c). The LR model demonstrated strong performance, achieving an area under the receiver operating characteristic curve (AUC) of 0.79. When the same set of features was used in a random forest (RF) model, the AUC increased to 0.83. Furthermore, the study evaluated the commercially available KidneyIntelX model, emphasizing the importance of external validation and model calibration for clinical deployment. The authors also recommended the development of more broadly applicable and interpretable AI models, underscoring the value of using routine laboratory data to enhance generalizability and avoid limitations due to test-specific requirements. The study also proposed integrating such models into electronic health record (EHR) systems to facilitate the effective transfer of knowledge into clinical decision-making processes.

Another study concentrated on patients with type 1 diabetes mellitus [7], using a diverse set of features—including albumin excretion rate (AER), estimated serum creatinine (eSCR), eGFR, HbA1c, duration of insulin-dependent diabetes mellitus (IDDM), age, and alcohol consumption—to evaluate twelve machine learning models, including LR, RF, XGBoost, and multilayer perceptron (MLP). In contrast to prior studies that primarily focused on predicting end-stage renal disease (ESRD), this work introduced an innovative approach to predict the overall ten-year risk of CKD development in type 1 diabetic patients. The authors proposed an advanced heterogeneous ensemble model, termed STK, which exhibited exceptional predictive performance: mean accuracy of 0.97, specificity of 0.98, sensitivity/recall of 0.96, precision of 0.98, F1-score of 0.97, Kappa and Matthews correlation coefficient (MCC) scores of 0.94, AUROC of 0.99, and precision–recall AUC of 0.99. This study successfully presented a novel and highly accurate method for predicting CKD onset risk in individuals with type 1 diabetes.

Compared to aforementioned studies, which primarily relied on laboratory data, our approach exclusively employs lifestyle features for risk prediction. This strategy not only reduces the cost associated with data acquisition but also facilitates patient self-monitoring, thereby lowering the frequency of clinical visits and alleviating the healthcare burden. Moreover, the predictive is no longer restricted to individuals with type 1 or type 2 diabetes.

Prediction based on laboratory data has reached a relatively mature stage, and in our study, we employed three widely used machine learning models that are well-suited for lifestyle tabular data—AutoGluon [8], TabPFN [9], and XGBoost [10]—and compared their predictive performance as well as their decision-making mechanisms. Among these, AutoGluon, an automated machine learning (AutoML) framework, demonstrated particularly strong predictive accuracy when applied to lifestyle tabular data, yielding results that were closely aligned with clinical judgment. The main contributions of this work are as follows:

- **To the best of our knowledge, this is the first study to predict CKD risk solely from lifestyle-related features by systematically benchmarking popular tabular learning models.**
- **Providing a patient-oriented self-assessment tool for risk prediction, simultaneously reduces healthcare expenditures associated with laboratory measurements.**
- **Expanding the prediction scope beyond type 1 and type 2 diabetes patients, thereby increasing its potential public health applicability.**
- **The design of GUI enables seamless user interaction and promotes system accessibility.**

II. DATA AND METHODOLOGY

The methodology of this study comprises six main stages. The first stage is Exploratory Data Analysis (EDA), aimed at understanding the dataset, particularly its lifestyle-related tabular features and classification structure. The second stage involves data preprocessing, where raw data are cleaned and transformed for modeling readiness. The third and fourth stage focuses on model training and evaluation performances, identifying the most suitable model. The fifth stage will utilize SHAP analysis to check the features' clinical credibility; if the results are inconsistent, an alternative model will be applied. At the final stage, the best-performing model is integrated into a GUI for simpler manipulation. The overall workflow is shown in Fig. 1.

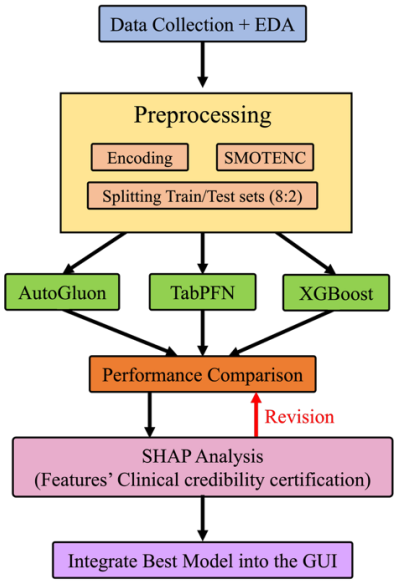


Fig. 1. Block Diagram of the Proposed Method

A. Data Collection, Description, and Exploratory data analysis

The dataset utilized in this study was obtained from BIRDEM General Hospital in Dhaka, Bangladesh, under the supervision of Dr. Rafi Nazrul from the Department of Nephrology, and was approved by the Ethical Review Committee of the Diabetic Association of Bangladesh [11]. The dataset consists of 4,000 records, covering 400 DM patients and includes 22 features such as gender, occupational type, family history of diabetes, and data collected over the past 10 years on BMI, comorbidities, lifestyle factors (including diet, physical activity, smoking, and water intake), urinary status, urinary infection and daily caloric intake. Among the variables, Diabetic Year, Age, BMI, and Calorie Intake were treated as numerical features, while the others were categorical, as presented in TABLE I. The dataset is highly imbalanced, with 386 cases of CKD and 3,614 non-CKD cases out of 4,000 records.

TABLE I. FEATURES IN THE DATASET AND THEIR DESCRIPTION

Feature	**description**	**Statistics**	
		Mean ± std	***variance***
Job	1-normal, 2-intermediate, 3-heavy	1.28 ± 0.60	0.36
Family Background of Diabetes	1-yes, 0-no	0.42 ± 0.49	0.24
Diabetic Year	Years since first diagnosis	9.27 ± 6.48	42.04
Age	Years	50.88 ± 10.95	119.95
BMI	kg/m^2	24.59 ± 3.77	14.24
Follow Suggested Diet	1-yes, 0-no	0.77 ± 0.42	0.18
Regularly Take Oral Anti-diabetic Medicine	1-yes, 0-no	0.75 ± 0.43	0.19
Take Insulin	1-yes, 0-no	0.55 ± 0.50	0.25
Hypertension	1-yes, 1-no	0.29 ± 0.45	0.21
Heart Disease	1-yes, 2-no	0.24 ± 0.43	0.18
Sleep	1-7 to 9 hours of sleep, 0-abnormal sleep time	0.59 ± 0.49	0.24
Water Consumption	1-more than 2 liters/day, 0-less than 2 liters/day	0.91 ± 0.29	0.08
Smoke	1-yes, 0-no	0.14 ± 0.34	0.12
Walk Regularly	1-more than 30 minutes per day, 0-less than 30 minutes per day	0.64 ± 0.48	0.23
Urination Properly	Urination without discomfort, obstruction, or medical issues, (1-yes, 0-no)	0.85 ± 0.36	0.13
Urinary Infection	1-yes, 0-no	0.24 ± 0.43	0.18
Taking Pain Killer	1-yes, 1-no	0.29 ± 0.46	0.21
Calorie Intake	Daily calories intake	1453.70 ± 211.55	44754.5

B. Data Preprocessing

The data preprocessing pipeline in this study involved several steps to ensure model compatibility and enhance learning performance. First, the categorical feature "Gender" was encoded using binary values, with female assigned as 0 and male as 1. Second, the features "Zarda use" and "Betel Leaf use" were removed, as such substances are not commonly consumed in Taiwan and were deemed irrelevant to the study objectives. Additionally, variables including height, weight, and average weight were excluded due to their high correlation with BMI, which was already presented in the dataset. This step was taken to prevent multicollinearity. In terms of data transformation, 16 categorical variables were processed through label encoding to ensure the machine learning models could interpret categorical information effectively. Meanwhile, the three continuous features were standardized to a uniform scale to prevent model bias caused by differences in magnitude. For dataset partitioning, the data was split into training and testing sets using an 80:20 ratio. However, given the highly imbalanced nature of the dataset—with a predominance of categorical features and only three continuous variables—Synthetic Minority Over-sampling Technique for Nominal and Continuous data (SMOTENC) [12] was employed for sample augmentation. SMOTENC generates synthetic minority class samples by considering both numerical and categorical variables, thereby alleviating class imbalance while preserving feature structure. This approach ensured consistency in the distribution of both training and testing sets, ultimately enhancing the accuracy and robustness of the model during training.

C. *Experiment*

In this study, we employed AutoGluon, developed by Amazon Web Services, to conduct model training [13]. AutoGluon simplifies the processes of model selection, training, and deployment, thereby improving the efficiency and usability of the machine learning pipeline. Additionally, it integrates stacking and bagging strategies to enable robust ensemble learning, enhancing predictive performance while maintaining computational efficiency.

In addition to AutoGluon, we independently trained two widely used machine learning models for comparative analysis. TabPFN, a transformer-based probabilistic model, is designed to process entire tabular datasets end-to-end during both training and inference. It demonstrates strong adaptability to structured tabular data, which is particularly well-suited to our case, as our dataset consists of lifestyle-related tabular features rather than continuous laboratory measurements commonly used in conventional medical AI.

On the other hand, XGBoost is a gradient-boosted decision tree model that is naturally well-suited for binary classification tasks and aligns well with the binary nature of most of our input features. It has been extensively applied in healthcare risk prediction and is frequently adopted as a baseline for evaluating emerging models. Therefore, TabPFN and XGBoost were selected as baseline models for performance comparison against AutoGluon. A detailed comparison of the predictive performance across all models is provided in the following section.

D. *Integration Best Model to GUI*

Fig. 2 illustrates the main functions of the GUI developed in this study. The left panel serves as the input area, where diabetic patients can enter relevant feature values and obtain model predictions by clicking the "Run test" button. The right panel displays two types of visual information: (1) the predicted probability of CKD occurrence for the current test, and (2) a probability trend chart of historical diagnostic results. Different risk levels are highlighted with color coding to facilitate interpretation by clinical staff and to support subsequent treatment decisions.

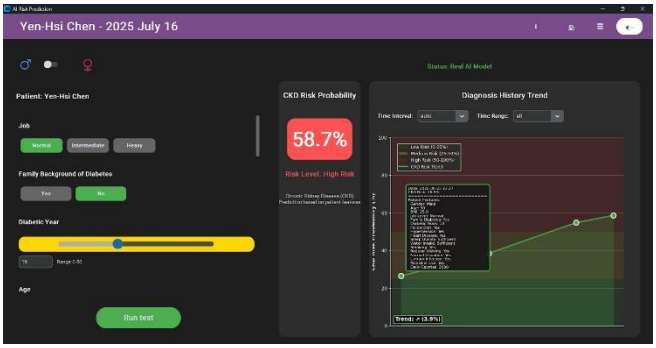


Fig. 2. GUI designed for AI-Based CKD. Input and Execution Result page

III. RESULT AND DISCUSSION

Table II and Fig. 3 present the comparative analysis of multiple machine learning models applied to CKD prediction, evaluated using five key performance metrics. The results indicate that TabPFN achieved the highest scores in ROC AUC (0.9950), F1-score (0.9642), and Recall (0.9617), demonstrating its strong capability in identifying true positive cases. In contrast, AutoGluon outperformed other models in terms of Accuracy (0.9813) and Precision (0.9755), reflecting its effectiveness in minimizing false positives while maintaining overall classification accuracy.

TABLE II. MODEL PERFORMANCE COMPARISON

	AutoGluon	TabPFN	XGBoost
Accuracy	0.9813	0.9648	0.8638
ROC AUC	0.9778	0.995	0.9436
F1_Score	0.8948	0.9642	0.8718
Precision	0.9755	0.9666	0.8231
Recall	0.8264	0.9617	0.9267

Overall, both AutoGluon and TabPFN exhibited excellent and consistent performance across all evaluation metrics, highlighting their strengths in handling classification tasks in the risk-prediction domain. By comparison, although XGBoost also demonstrated reasonable predictive power, its performance varied more noticeably across different metrics, indicating relatively lower stability.

In summary, TabPFN and AutoGluon demonstrated superior and stable predictive capabilities for assessing CKD development risk among diabetic patients. These models show strong potential for further development and implementation in future research and clinical applications.

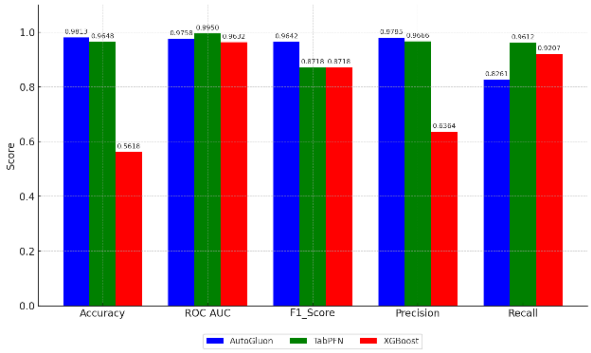


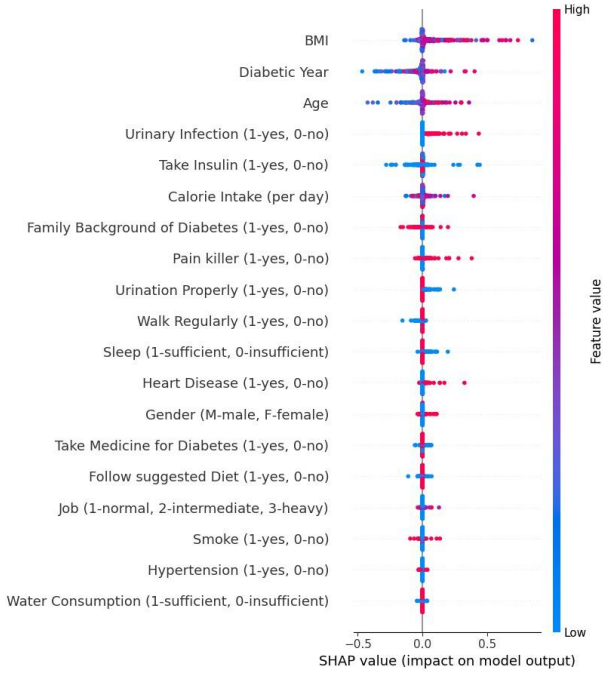
Fig. 3. Model Performance of AutoGluon, TabPFN, and XGBoost

Despite the outstanding predictive performance demonstrated by both TabPFN and AutoGluon, a critical limitation remains in their lack of interpretability. These models are often regarded as "black boxes" due to their reliance on complex feature interactions and data correlations rather than on transparent and intuitive rules. This opacity may undermine trust and adoption in clinical settings, where understanding the reasoning behind model predictions is essential for effective decision-making.

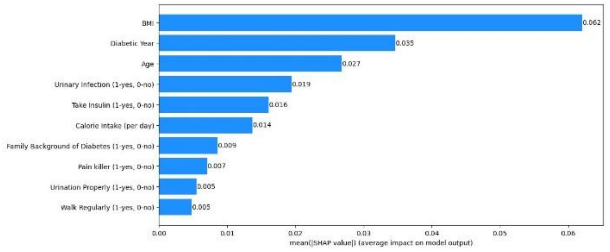
To address this limitation, we employed Shapley Additive Explanations (SHAP) [14], a method grounded in cooperative game theory, to fairly attribute the contribution of each input feature to the model's output. Through SHAP analysis, we were able to quantify and interpret the influence of individual features on CKD risk predictions, thereby enhancing the transparency of the model's decision-making process and aligning it with the expectations of clinical practitioners.

This interpretability mechanism enables clinicians and experts to cross-validate model outputs with their domain expertise, improving the model's credibility and acceptability in real-world medical applications.

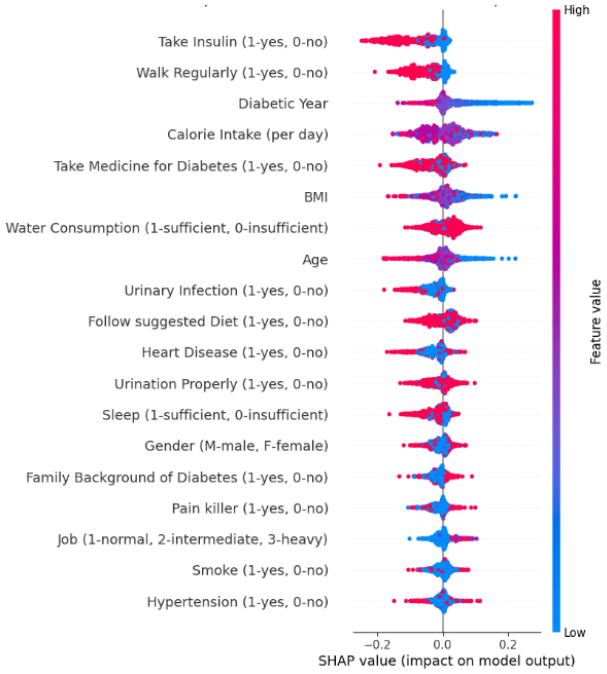
SHAP value plots of the AutoGluon model represents that the most important feature was BMI (Fig. 4(a)). A higher BMI generally indicates excess body weight, which imposes a greater metabolic burden and consequently elevates the risk of developing CKD [15]. The second priority feature was Diabetic Year. A longer disease duration implies prolonged exposure to hyperglycemia, thereby increasing the risk of renal impairment [16]. The third priority feature was Age. As patients grow older, physiological decline increases their susceptibility to kidney disease [17]. The fourth feature is Urinary Infection [18]. Pyelonephritis caused by bacterial infection can result in inflammation and fibrosis of the renal parenchyma, ultimately leading to a progressive decline in renal function, and the risk of infection is particularly elevated among patients with diabetes [19]. The relative contribution of each feature has been shown in Fig. 4(b). Generally, AutoGluon achieved robust performance across key clinical indicators, underscoring its potential value for early prevention.



(a) AutoGluon SHAP splot



(b) AutoGluon's feature contribution bar plot



(c) TabPFN SHAP plot

Fig. 4. SHAP summary plot

It is noteworthy that although TabPFN achieved outstanding overall predictive performance, SHAP analysis revealed that its feature importance rankings were often inconsistent with established clinical experience (Fig. 4(c)). Such discrepancies are difficult to accept in medical contexts, thereby limiting the model's clinical credibility. In contrast, AutoGluon not only demonstrated strong predictive accuracy but also produced feature attribution results that aligned closely with clinical knowledge. This consistency underscores its greater applicability in predicting and preventing CKD risk among diabetic patients, positioning AutoGluon as the most suitable model in this study.

Overall, the SHAP summary plots highlighted the relative importance and interactions of individual features during AutoGluon model training. These findings are consistent with existing literature on the associations between clinical features, diabetes, and its complications, thereby further validating the robustness and reliability of the model's predictions and supporting its feasibility for clinical application.

IV. Conclusion

In this study, we proposed a machine learning–driven framework for predicting CKD risk in diabetic patients using lifestyle features as the primary input variables. Among the models evaluated, AutoGluon consistently outperformed other approaches in terms of predictive accuracy and precision, while also providing clinically coherent feature attributions through SHAP analysis. This combination of robust predictive capability and transparent interpretability highlights AutoGluon's potential as a trustworthy and clinically meaningful decision support tool.

By emphasizing lifestyle-related features as the core predictors, our framework departs from the conventional reliance on laboratory-based indicators and demonstrates the feasibility of a non-invasive, cost-effective, and patient-centered approach to risk stratification. This innovation not

only reduces healthcare costs and alleviates the demand on clinical resources but also aligns with the growing emphasis on preventive medicine, precision health strategies, and patient empowerment.

Looking forward, future work will involve expanding training to larger, more demographically diverse cohorts, integrating complementary clinical and biochemical markers, and performing rigorous multi-center external validation to establish real-world applicability. Beyond technical improvements, embedding this framework into digital health ecosystems and mobile health platforms offers an opportunity to bridge the gap between predictive analytics and everyday clinical practice, thereby fostering equitable access to advanced decision-support tools across healthcare settings.

Ultimately, this research advances the field of explainable artificial intelligence in medicine by providing both methodological novelty and clinically relevant impact. It contributes to the paradigm shift toward proactive disease prevention, precision medicine, and patient engagement, while also offering valuable insights for healthcare policy and resource optimization in the management of diabetes-associated CKD.

REFERENCES

[1] Chen, T. K., Knicely, D. H., & Grams, M. E. (2019). Chronic kidney disease diagnosis and management. *JAMA*, *322*(13), 1294.

[2] Alicic, R. Z., Rooney, M. T., & Tuttle, K. R. (2017a). Diabetic kidney disease. *Clinical Journal of the American Society of Nephrology*, *12*(12), 2032–2045.

[3] Wu, I., Wu, M., Barone, S., Cabrera, C., Sanchez, J. J., Retat, L., Mitchyn, M., & Wu, M. (2025a). Projection of the prevalence and economic burden of chronic kidney disease in Taiwan from 2022 to 2027 (inside ckd): A microsimulation study. *Nephrology*, *30*(5).

[4] Ferguson, T., Ravani, P., Sood, M. M., Clarke, A., Komenda, P., Rigatto, C., & Tangri, N. (2022). Development and external validation of a machine learning model for progression of CKD. *Kidney International Reports*, *7*(8), 1772–1781.

[5] Zhu, H., Qiao, S., Zhao, D., Wang, K., Wang, B., Niu, Y., Shang, S., Dong, Z., Zhang, W., Zheng, Y., & Chen, X. (2024). Machine learning model for cardiovascular disease prediction in patients with chronic kidney disease. *Frontiers in Endocrinology*, *15*.

[6] Tangri, N., & Ferguson, T. W. (2022). Artificial Intelligence in the identification, management, and follow-up of CKD. *Kidney360*, *3*(3), 554–556.

[7] Chowdhury, M. N., Reaz, M. B., Ali, S. H., Crespo, M. L., Cicuttin, A., Ahmad, S., Haque, F., Bakar, A. A., Razak, M. I., & Bhuiyan, M. A. (2024). Machine learning algorithms for predicting the risk of chronic kidney disease in type 1 diabetes patients: A retrospective longitudinal study. *Neural Computing and Applications*, *36*(26), 16545–16565.

[8] Erickson, N., Smola, A., Li, M., Larroy, P., Zhang, H., Shirkov, A., & Mueller, J. (2020). AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. *arXiv*.

[9] Hollmann, N., Müller, S., Eggensperger, K., & Hutter, F. (2022). TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second. *arXiv*.

[10] Chen, T., & Guestrin, C. (2016). XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[11] Anan, A. (2025a, July 2). A dataset on demographic and lifestyle factors for prognosticating chronic kidney disease progression in diabetic patients. Mendeley Data.

[12] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357.

[13] Huang, C.-H., Tang, L.-C., Chen, Y.-C., Lin, C.-H., Liang, C.-H., Hsiao, E., & Mai, F.-D. (2025a). Enhancing diabetes diagnosis and complications prediction with automated machine learning in a clinical decision support system. *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 0778–0783.

[14] Lundberg, S., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *arXiv*.

[15] Kim, B.-Y., Choi, D.-H., Jung, C.-H., Mok, J.-O., & Kim, C.-H. (2021). Associations between obesity, weight change and decreased renal function in Korean type 2 diabetic patients: A longitudinal follow-up study. *BMC Endocrine Disorders*, *21*(1).

[16] Kim, C. S., Suh, S. H., Choi, H. S., Bae, E. H., Ma, S. K., Kim, B., Han, K.-D., & Kim, S. W. (2025). Impact of diabetes duration and hyperglycemia on the progression of diabetic kidney disease: Insights from the KNHANES 2019-2021. *World Journal of Diabetes*, *16*(5).

[17] Fabre, L., & Bevilaqua Rangel, É. (2024). Age-related markers and predictors of diabetic kidney disease progression in type 2 diabetes patients: A retrospective cohort study. *Therapeutic Advances in Endocrinology and Metabolism*, *15*.

[18] Belyayeva, M. (2024, February 28). Acute Pyelonephritis. StatPearls [Internet].

[19] Zasloff, M. (2018). Why are diabetics prone to kidney infections? Journal of Clinical Investigation, 128(12), 5213–5215.