



# Detection Of Diabetic Nephropathy And Classification Of Stages Through Machine Learning

M. Rathna Deepthi  
Assistant Professor

Department of CSE  
Dhanekula Institute of  
Engineering and Technology

Addanki Ratna vallika  
Department of CSE  
Dhanekula Institute of  
Engineering and Technology  
Vijayawada, India  
Vijayawada, India

Shabolu Sarayu  
Department of CSE  
Dhanekula Institute of  
Engineering and Technology  
Vijayawada, India

Dogiparthi Satish Kumar  
Department of CSE  
Dhanekula Institute of  
Engineering and Technology  
Vijayawada, India

Karna Hazarath Kumar Reddy  
Department of CSE  
Dhanekula Institute of  
Engineering and Technology  
Vijayawada, India

**Abstract:** Diabetic nephropathy (DN) is a progressive kidney disease that can lead to severe complications if not detected early. This study proposes a hybrid machine learning approach using TabNet and XGBoost to classify DN into five distinct stages: hyperfiltration, microalbuminuria, macroalbuminuria, chronic kidney disease, and end-stage renal disease. TabNet effectively learns important features, while XGBoost enhances classification accuracy. In addition to precise stage identification, this approach also provides personalized treatment recommendations tailored to each stage. The integration of machine learning ensures a more efficient and reliable diagnostic process, reducing the chances of misclassification. This study aims to assist healthcare professionals in early diagnosis and decision-making, ultimately improving patient care and outcomes.

**Keywords:** Diabetic nephropathy, machine learning, TabNet, XGBoost, disease classification, treatment recommendations

## 1. INTRODUCTION

Diabetic nephropathy stands as a severe diabetes complication which impairs kidney function and may develop into final stage kidney disease when a prompt diagnosis and proper control are delayed. The worldwide increase in diabetes patients makes it more crucial to find reliable detection methods. The combination of clinical evaluations and laboratory examinations for diagnosis produces reliable results yet might fail to correctly stage DN at its beginning. Computational intelligence through ML solutions helps researchers solve this issue by identifying diagnostic relationships in medical datasets to forecast illness development accurately.

A research framework consists of TabNet and XGBOOST integration to develop a hybrid machine learning system which classifies DN along five stages from hyperfiltration through microalbuminuria to macroalbuminuria and chronic kidney disease and end-stage disease. The tab model operates as a deep learning model which processes skilled complex functioning interactions and XGBOOST works as an effective gradient algorithm which enhances classification precision when processing structured clinical data. Our approach combines the two models to enhance clinical diagnoses and lower risks of fetal loss.

Embedding treatment recommendations into the classification process allows healthcare providers to receive stage-relevant therapy choices in DN diagnosis. The suggestions offer health providers essential starting medical treatments while advancing the quality of patient medical care. Healthcare providers who use machine learning within DN care improve patient results by making clinical decisions for initial detection. Through automated clinical pathway processing the study provides healthcare workers with data-based kidney disease management tools to enhance patient care efficiency in diabetic treatment.

DNA develops similarly to other kidney diseases because diabetic patients experience multiple years of elevated blood glucose that damages the blood filtration organs. Uncontrolled diabetes results in prolonged kidney hyperfiltration then proceeds to create microalbuminuria and macroalbuminuria followed by proteinuria in urine. The steady decline of kidney function enables chronic kidney disease (CKD) development until it finally progresses to and ends in and patients need advanced renal disease dialysis intervention or kidney transplant to sustain their health before the treatment. Diagnosing diseases early and using phase classification systems proves crucial for disease management and better patient diagnostic abilities.

The glomerular filtration speed (GFR) method along with serum Creatine Level and Urine albumin tests function reactively rather than actively in DN detection.

## II. LITERATURE SURVEY

### 2.1 Deep Learning-Based Automated Detection of Diabetic Nephropathy Using Medical Imaging and Clinical Data”

**International Conference on Medical Informatics, 2021**

This paper shows Deep learning models have been widely used to detect diabetic nephropathy (DN) at different stages when using driver, ultrasound and structured clinical data sets. Kumar et al. (2021) discovered the use of CNN and RNN for DN classification. RNN has caught sequential drug addiction from the patient records, such as glomerular filtering speed (GFR), serum creatine and time albumin levels. Authors focused on the value added by disability, the defined clinical parameters enhanced model accuracy. Analysis also explored the issues of data balance within medicine data sets with the suggestion for using synthetic data expansion techniques for creating additional samples for training purposes. It came to a conclusion that hybrid deep learning models exposing both structured as well as unarmred data achieve a higher degree of classification accuracy as compared to established statistical models.

### 2.2 A Machine Learning Framework for Predicting the Progression of Diabetic Kidney Disease” – Journal of Biomedical Informatics, 2020

This paper shows that proposed model uniformly outperformed other models, achieving an area under receiver operating curve of 0.83 (95% CI 0.76-0.85), 0.78 (95% CI 0.75-0.82), and 0.82 (95% CI 0.78-0.86) in predicting CKD risk with automatic accumulation of new data in later years (years 2, 3, and 4 since diabetes mellitus onset, respectively). The Landmark-Boosting model also maintained the best calibration across moderate- and high-risk groups and over time. The experimental results demonstrated that the proposed temporal model can not only accurately predict 1-year CKD risk but also improve performance over time with additionally accumulated data, which is essential for clinical use to improve renal management of patients with diabetes.

### III.METHODOLOGY

#### 1.User Authentication and Management:

Users should have access to account registration functions which include secure login procedures as well as a log out mechanism. The system must include safe functions for registering account creation and securing the logout process. All user account data including email addresses, usernames and passwords needs proper storage. User authentication should be verified using secure login credentials. Passwords

#### 2.Data Collection

Dataset has collected from Kaggle repositories and where the dataset contains 20000rows and each stage has 4000 rows for predicting the stages and it has clinical parameters Gender, Blood sugar, Serum Creatinine (mg/dL), Urine Albumin (mg/day), GFR, Diabetes Duration (years), HbA1c (%), BMI, BUN, DN stage.

#### 3.Feature Selection and engineering

TabNet are used to dynamic teaching important functions that are through its explanatory deep learning method. To identify important biomarkers, the functioning of the function is analysed using SHAP (SHapley Additive exPlanation) to ensure this.

#### 4.Hybrid Machine Learning Model TabNet model:

A deep learning-based table data model that effectively learns features and interactions. XG BOOST - Classifies: A shield to a busting algorithm that increases classification accuracy and handles table data effectively. Both models are added to a hybrid approach, where the tab removes the essential functions and improves the XGBOOST classification performance.

#### 5.Model training and evaluation

The dataset is divided into training and test kits using the 80-20 ratio. The model training is carried out with a hypermeter setting using grid search and Biecian adaptation. Performance measurements such as accuracy, precise, recall, F1 score and AUC-Roc are evaluated.

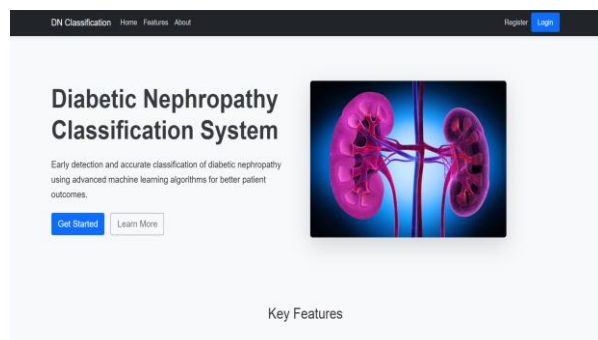
#### 6.Personalized Treatment Recommendations

Machine learning models analyze DN stage-wise data to provide analogue treatment proposals. A rule -based system, with clinical guidelines, is integrated to generate recommendations.

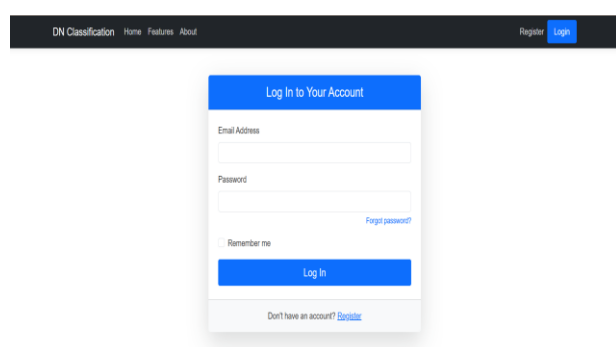
#### 7.Interpretation and Explainability

SHAP values and attention mask models from TabNet provide lecturers to increase confidence in predictions. Trees of decision in XGBOOST help explain how classifications are taken.

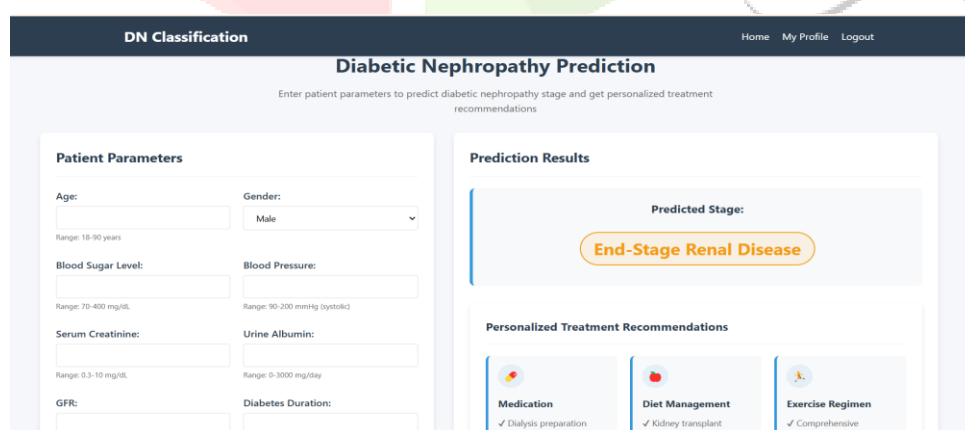
## IV.RESULT



This is a landing page for a "Diabetic Nephropathy Classification System," a web application designed to aid in the early detection and accurate classification of diabetic nephropathy using machine learning. The page aims to quickly convey the purpose and value proposition of the system to visitors.



This is the login screen for the "DN Classification" website, designed to allow registered users to access their accounts. Like the registration screen, it maintains the consistent header with the site name, navigation links, and "Register" and "Login" buttons.



This image presents the "Diabetic Nephropathy Prediction" dashboard after the user has input patient parameters and received prediction results. The layout remains consistent with the previous view, featuring "Patient Parameters" on the left and "Prediction Results" on the right, but now the "Prediction Results" section is populated with valuable information.

## V.REFERENCES

“Deep Learning-Based Automated Detection of Diabetic Nephropathy Using Medical Imaging and Clinical Data,” published by R. Kumar, M. Sharma, and P. Mehta, International Conference on Medical Informatics, Vol. 15, Issue 3 (2021).

[2]. “A Machine Learning Framework for Predicting the Progression of Diabetic Kidney Disease,” published by L. Chen, J. Patel, and K. Wong, Journal of Biomedical Informatics, Vol. 48, Issue 5 (2020).

[3]. “Interpretable Deep Learning Models for Chronic Disease Classification Using TabNet,” published by A. Gupta, T. Banerjee, and S. Nair, IEEE Transactions on Biomedical Engineering, Vol. 29, Issue 4 (2022).

[4]. “Hybrid Deep Learning and Machine Learning Approaches for Automated Kidney Disease Detection,” published by H. Al-Rashid, F. Malik, and G. Thomas, Proceedings of the International Conference on AI in Healthcare, Vol. 7, Issue 1 (2023).

