Comparison of artificial intelligence algorithms and expert approach in risk classification in nephrology

Dawid Pawuś[1,*], Szczepan Paszkiel[1] and Tomasz Porażko[2,3]

[1] *University of Technology, Prószkowska 76 Street, 45-758 Opole, Faculty of Electrical Engineering, Automatic Control and Informatics, Poland*

[2] *Departament of Internal Medicine and Nephrology, Institute of Medical Sciences, University of Opole, Oleska 48 Street, Opole, 45–052, Poland*

[3] *Departament of Nephrology with Dialysis Unit, Opole University Hospital, Witosa 26 Street, Opole, 46-020, Poland*

Abstract

This study presents a comparative analysis of various artificial intelligence (AI) and machine learning (ML) algorithms for risk classification in idiopathic membranous nephropathy (IMN), a complex kidney disease. The research evaluates seven different models, including K-Nearest Neighbors, Decision Trees, Random Forests, Support Vector Machines, Adaptive Boosting, LightGBM, and Multi-layer Perceptron. The results reveal that ensemble methods, particularly Random Forests, achieve the highest precision in classifying IMN risk levels, highlighting their potential in improving diagnostic accuracy and patient management in nephrology. The study underscores the importance of model selection and fine-tuning to optimize AI applications in clinical settings, providing a basis for future advancements in AI-driven nephrology.

Keywords

Artificial Intelligence, Machine Learning, Expert System, Classification, Kidney, Nephrology, Automation

1.Introduction

Nephrology is a critical area within medical sciences, with kidney diseases impacting approximately 850 million people worldwide, as indicated by recent registries [1]. Like other scientific disciplines, nephrology is currently experiencing a transformation driven by advanced technologies. A key component of this evolution is the close collaboration between IT engineering and medical professionals. This partnership is leading to an increase in research efforts and a growing interest in developing and implementing AI-driven systems, numerical and classification algorithms, as well as expert systems. These innovations have the potential to enhance, automate, and refine the processes of diagnosing, classifying, and ultimately improving the treatment outcomes for kidney diseases.

ITTAP'2024: 4th International Workshop on Information Technologies: Theoretical and Applied Problems, November 20–22, 2024, Ternopil, Ukraine, Opole, Poland *Corresponding author.
†These authors contributed equally. dawid.pawus@student.po.edu.pl (D. Pawuś);
s.paszkiel@po.edu.pl (S. Paszkiel);
tomasz.porazko@usk.opole.pl (T. Porażko)
0000-0003-3308-3474 (D. Pawuś); 0000-0002-4917-5712 (S. Paszkiel); 0000-0002-8516-2740 (T. Porażko)

Historically, nephrology has relied on manual analysis, which presents challenges when dealing with complex data sets, often leading to diagnostic errors and suboptimal treatment choices. The ongoing advancements in numerical algorithms and artificial intelligence offer new opportunities to enhance kidney health monitoring. This progress allows for greater diagnostic accuracy and more precise treatment strategies. The synergy between expert medical knowledge and the automation capabilities of algorithms is paving the way for systems that can effectively address complex nephrological challenges.

Glomerular diseases, a subset of kidney conditions, specifically target the glomeruli— tiny structures in the kidneys that filter blood and produce urine. These diseases can arise from various causes, including infections, autoimmune disorders, hypertension, diabetes, and other metabolic conditions.

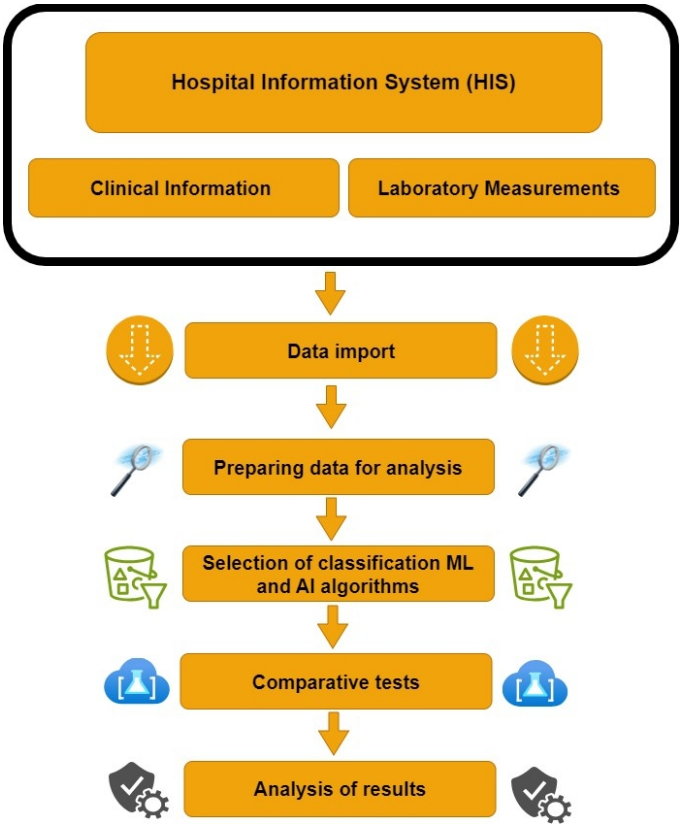

Figure 1: General outline of the procedure for research and verification of the problem.

This article provides a comparative approach to different AI algorithms that can be used in risk classification of patients with a nephrological disease called idiopathic membranous nephropathy (IMN). The concept is to create an extended and multi-task expert system that would be tasked with supporting nephrologists' decisions in their daily work in treating patients with kidney diseases. The risk classification module is therefore crucial from the perspective of the domain expert, who can determine the further course of action based on the patient's assignment to a specific group. The range of algorithms ranges from simple linear approaches to complex and advanced gradient solutions. The expertise of nephrologists is also leveraged along with advances in engineering. Focusing on specific applications, our goal is to show how these advances can impact the standards of patient care, ultimately improving the outcomes of

traditional approaches. Figure 1 provides a general outline of the methodology, scope of the study and the approach used to address the problem.

The next section will provide a detailed literature review and an overview of the state of the art.

2.State of the art and related works

This literature review examines the use of automation, artificial intelligence, and classification systems in the field of nephrology. It delves into how these advanced technologies streamline processes, increase diagnostic precision, and improve patient outcomes. Through a critical analysis of the current literature, the goal is to uncover trends and identify gaps, thereby guiding future research toward innovative solutions in the field of nephrology.

Advancements in artificial intelligence, innovation, and transformative technologies are pivotal in nephrology and dialysis [2–4]. Additionally, the future prospects of AI and modern technologies in managing kidney diseases are extensively reviewed [5, 6]. Research focusing on numerical models and machine learning techniques also holds considerable importance [7–9].

A key area in nephrology research is the integration of acute kidney injury (AKI) with AI-driven predictive algorithms [10, 11]. Another significant focus includes the development of predictive models in the context of personalized medicine [12]. AI-based clinical applications in nephrology are explored in works by [13] and [14]. Furthermore, ethical considerations surrounding AI applications are discussed in [15] and [16].

Recent research underscores the integration of AI and machine learning within nephrology. For instance, [17] develops short-term prognosis prediction models for severe AKI patients undergoing PIRRT (Prolonged Intermittent Renal Replacement Therapy), utilizing algorithms such as Naive Bayes and Random Forest, and highlighting the critical role of serum electrolyte management. Another study [18], introduces Sugeno's fuzzy inference system for regression problems involving numerous variables and limited data, demonstrating superior performance over traditional methods.

In addition, [19] presents health-disease phase diagrams (HDPDs) for visualizing disease onset probabilities, while [20] develops CKD.Net, a hybrid model that achieves high accuracy in predicting CKD stages. Other significant contributions include [21], which focuses on precise kidney volume measurement using AI, and [22], which addresses PRCC (Papillary Renal Cell Carcinoma) stage classification. Furthermore, [23] showcases the importance of machine learning in predicting AKI and diagnosing PDAC (Pancreatic Ductal Adenocarcinoma).

A study derived and validated an ML risk score for predicting diabetic kidney disease (DKD) progression using biomarker and electronic patient data [24]. Research into interpretable ML for early AKI prognosis emphasizes model interpretability benefits [25]. ML has also been utilized to predict primary nephrotic syndrome pathology without biopsy and identify hub genes in membranous nephropathy [26, 27]. A model for predicting idiopathic membranous nephropathy prognosis is discussed in [28]. ML models predicting rituximab underdosing in membranous nephropathy show high accuracy, sensitivity, and specificity [29]. Another study reports a machine learning framework for diagnosing membranous nephropathy from whole-slide images, achieving a 97.32% F1-score [30].

A belief rule-based system for diagnosing primary membranous nephropathy shows significant reliability with 98% sensitivity, 96.9% specificity, 97.8% accuracy, and an AUC of 0.93 [31]. A predictive model for long-term renal function impairment post-minimally invasive partial nephrectomy has a concordance index of 0.75 [32]. The use of AI and ML in dialysis is

reviewed, focusing on diagnostics, prognosis, and treatment optimization [33]. Guidelines for proper application and transparent reporting of ML models in biomedical research emphasize best practices [34]. Recent literature highlights the crucial role of clinical prediction models in modern healthcare, emphasizing challenges and the need for transparent reporting to assess their quality [35]. The literature underscores the importance of understanding diagnostic and prognostic prediction models, addressing issues such as model development, validation, and sample size considerations to improve clinical decision-making [36–38], as well as valuable research [39, 40].

3. Structure, research scope and methodology in the problem of automatic risk classification

This section explores the development of an expert system for diagnosing and treating idiopathic membranous nephropathy, focusing specifically on the risk classification module. The system is structured as a hierarchical decision tree that integrates clinical data, diagnostic criteria, and therapeutic guidelines to assist nephrologists in managing IMN.

3.1. General overview of system modules

The expert system consists of several key components, which are shown in Figure 2 and briefly described below:

1. Diagnostic Module – the diagnosis of (IMN) is confirmed through a combination of patient history, diagnostic tests, and kidney biopsy. The evaluation also considers the likelihood of progression to stage 5 chronic kidney disease (CKD5) by analyzing glomerular filtration rate, levels of proteinuria, and serum albumin concentrations.
2. Risk Stratification – patients are classified into four distinct risk categories – low, intermediate, high, and very high – based on factors such as estimated glomerular filtration rate (eGFR) and proteinuria levels. Each category is associated with specific management strategies tailored to the risk profile.
3. Therapeutic Module – treatment recommendations vary from conservative management to more aggressive immunosuppressive therapies, depending on the risk category. Options include ACE inhibitors or ARBs, rituximab, calcineurin inhibitors, cyclophosphamide, and corticosteroids.
4. Treatment Effectiveness Assessment Module and Treatment Continuation Module – patient response to treatment is re-evaluated after six months. Based on clinical indicators, the treatment plan may be continued, modified, or altered. Protocols are provided for deciding on the appropriate course of action.
5. Follow-up Module – guidelines are provided for the long-term monitoring and management of patients who achieve partial or complete remission, including strategies for addressing treatment resistance and managing relapses.
6. User Interface – the system includes a user-friendly graphical interface designed for ease of use by healthcare professionals.

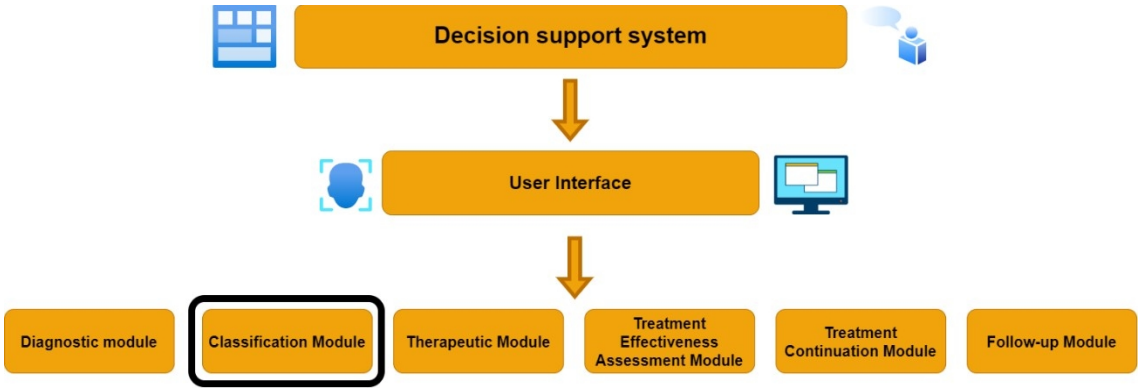

Figure 2: General diagram of the system with the module being the object of the study marked.

In this article, we focus solely on the Risk Stratification Module (see Fig. 2) of the expert system, which plays a crucial role in assigning patients to specific risk groups for targeted management. The comparative analysis involves various AI algorithms, from basic linear methods to advanced gradient-based solutions, alongside traditional expert approaches.
The details will be explained in the following sections of the article.

The remaining components of the system – such as diagnostic, therapeutic, and followup modules – will be explored in future research. This initial focus on risk classification aims to demonstrate how AI advancements can enhance traditional nephrological practices and improve patient outcomes.

3.2. Input-output model of the classification system

This subsection details the process of modeling the classification system, presenting the input variables and their corresponding output classifications. The classification model processes clinical parameters to categorize IMN patients into risk groups using a Multi Input Single Output (MISO) framework. The authors used the information provided in the guidelines in [41]. The 11 inputs to this model, detailed below, include a range of variables:

1. eGFR (Estimated Glomerular Filtration Rate) – a measure of kidney function estimating how well the kidneys filter blood.
2. Proteinuria – the level of protein present in the urine.
3. Serum Albumin Concentration – the concentration of albumin in the blood.
4. Response to ACEi/ARB Treatment – the percentage reduction in proteinuria after 6 months of treatment with ACE inhibitors or ARBs.
5. Serum anti-PLA2R Concentration – the level of anti-PLA2R antibodies in the serum.
6. Urinary α1-microglobulin Concentration – the concentration of α1-microglobulin in the urine.
7. Urinary IgG Concentration – the concentration of IgG in the urine.
8. Urinary β2-microglobulin Concentration – the concentration of β2-microglobulin in the urine.
9. Selectivity Index – the ratio of different urinary protein components indicating the selectivity of proteinuria.

10. Nephrotic Syndrome Symptoms – the presence of severe symptoms associated with nephrotic syndrome (binary value).
11. Rapid Renal Function Impairment – a swift decline in kidney function not attributable to other diseases (binary value).

The classification model uses machine learning to provide accurate and reliable risk assessments. Risk categories are defined by specific clinical criteria, and integrating machine learning techniques significantly increases the model's capabilities:

1. Low Risk
- eGFR > 60 ml/min/1.73 m^2
- Proteinuria < 3.5 g/d
- Serum albumin > 30 g/l OR
- eGFR > 60 ml/min/1.73 m^2
- Proteinuria < 3.5 g/d or a reduction > 50% after 6 months of ACEi/ARB treatment

2. Moderate Risk
- eGFR > 60 ml/min/1.73 m^2
- Proteinuria > 3.5 g/d and no reduction > 50% after 6 months of ACEi/ARB treatment OR
- Does not meet high risk criteria

3. High Risk
- eGFR < 60 ml/min/1.73 m^2 and/or Proteinuria > 8 g/d for 6 months OR
- eGFR > 60 ml/min/1.73 m^2
- Proteinuria > 3.5 g/d and no reduction > 50% after 6 months of ACEi/ARB treatment, plus one of the following:
    a. Serum albumin < 2.5 g/dl
    b. Serum anti-PLA2R > 50 RU/ml
    c. Urinary β2-microglobulin > 40 μg/min
    d. Urinary IgG > 1 μg/min
    e. Urinary β2-microglobulin > 250 mg/d
    f. Selectivity Index > 0.20
    g. Very High Risk
    h. Life-threatening nephrotic syndrome symptoms OR
- Rapidly progressive renal impairment not caused by other diseases.

Machine learning algorithms are excellent at identifying complex interactions between multiple input variables that may not be apparent with rule-based logic alone. This leads to a more nuanced understanding of a patient's risk profile. The guidelines above provide a framework for categorizing data, but they cannot account for every possible combination of lab results and other factors. Nor could a rule-based system easily provide a way to determine which output category values fall into. This gives AI and ML methods an advantage in these types of tasks.

Machine learning models leverage historical patient data to identify patterns and correlations that enhance the precision of risk classification predictions. This approach, driven by data, refines decision boundaries and reduces the likelihood of misclassification. A key benefit of

machine learning models is their capacity to adapt and become more accurate as new data is introduced. Furthermore, certain machine learning models offer insights into the probability of data belonging to specific classes.

The next section will describe in detail the considered algorithms.

4.Overview of algorithms used in the study

In any problem where automatic classification of collected data is needed, selecting the right AI models and numerical algorithms is a key step in developing a robust classification system. Seven AI and ML-based models were selected and tested to provide the best solution to the task. The models feature different algorithms used, such as tree-based methods, ensemble techniques, support vector machines, and neural networks. By using different methods, a comprehensive evaluation and comparison is provided. It also provides researchers with an answer as to which model performs best in classifying the risk of IMN patients. Below is a brief description of each algorithm:

1. **K-Nearest Neighbors (KNN)** is a straightforward, non-parametric method used for both classification and regression. It classifies a data point by considering the most common class among its k nearest neighbors in the feature space. While KNN is simple to grasp and implement, it can become computationally heavy with larger datasets.
2. **Decision trees (DT)** offer a flexible approach to machine learning for both classification and regression tasks. They partition data into subsets based on feature values, creating a tree-like structure of decisions. Though easy to interpret, decision trees can overfit the data if not properly managed through pruning.
3. **Random forests (RF)** are an ensemble learning method that builds and combines multiple decision trees to achieve more accurate and stable predictions. By using different subsets of data and features to construct each tree, random forests enhance generalization and help reduce overfitting.
4. **Support Vector Machines (SVMs)** are effective classifiers that determine the optimal hyperplane for separating classes in the feature space. They work well in high-dimensional settings and can address both linear and non-linear classification problems through kernel functions.
5. **Adaptive Boosting (AdaBoost)** is an ensemble technique that merges several weak classifiers into a strong one. It adjusts the weights of misclassified instances during each iteration, placing greater emphasis on harder-to-classify examples. Despite its simplicity, AdaBoost can be sensitive to noisy data.
6. **LightGBM** is a gradient boosting framework designed for efficiency and scalability, particularly with large datasets. It uses tree-based learning algorithms and includes optimizations to improve both speed and memory efficiency.
7. The **Multi-layer Perceptron (MLP)** Classifier is a type of neural network with multiple layers of neurons. It is adept at capturing complex patterns in data and can be used for classification and regression. However, MLPs require careful tuning of hyperparameters and can involve lengthy training processes.

This chapter briefly introduces seven different AI and ML models for their usefulness in classifying the risk of IMN patients. The models reviewed include K-Nearest Neighbors,

Decision Trees, Random Forests, Support Vector Machines, Adaptive Boosting, LightGBM, and the Multi-layer Perceptron Classifier. Each model offers unique strengths: KNN is straightforward but computationally intense, DTs are interpretable but prone to overfitting. RFs improve stability and generalization, SVMs excel in high-dimensional spaces and AdaBoost focuses on hard-to-classify examples but can be sensitive to noise. LightGBM provides efficiency and scalability for large datasets, and MLPs capture complex patterns but require extensive tuning.

The next chapter will detail the training process of these models and provide a discussion of the results obtained from their application to the classification task.

5.Training of models and discussion of achieved results

In this section, we present the insights gained from the analysis and the results of classifying nephrology patient data using various machine learning models described in Section 4.

The original dataset included laboratory results relevant to IMN risk classification. To mitigate class imbalance and improve model performance, synthetic data augmentation was applied, generating new data points that mirror the distribution and features of the existing data. For model training, 200 datasets were used, evenly distributed among the four risk categories: low, medium, high, and very high, with 50 datasets in each category. An identical number of datasets were allocated for testing. This approach ensured that the models had ample data for training and a reliable set for evaluation.

The comparison of precision scores (see Fig. 3) across various machine learning models in the context of nephrology risk classification reveals important insights into their performance on the dataset prepared and described in the preceding sections.

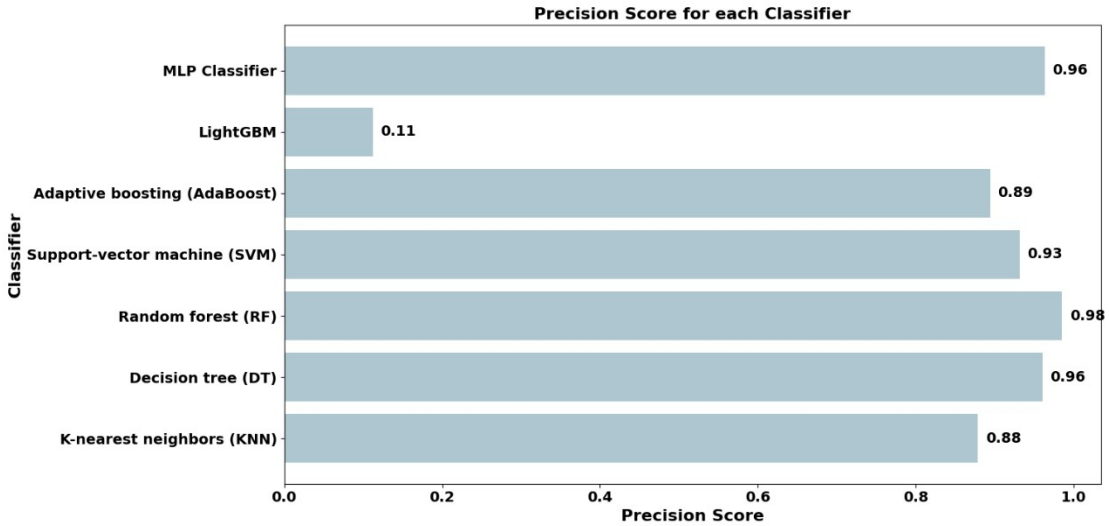

Figure 3: Precision results for tested models.

By analyzing the results in Figure 3, we can discuss in detail the results and implications of each classifier:

1. MLP Classifier – the MLP classifier achieved a high precision score of 0.96, indicating its robustness in accurately predicting the positive class in the IMN risk classification task. This model's architecture, which simulates the neural networks found in biological brains, has proven effective in capturing the complex, non-linear relationships inherent in the clinical dataset.
2. LightGBM – in contrast, the LightGBM model exhibited a notably lower precision score of 0.11. This result is surprising given LightGBM's reputation for efficiency and high performance in many classification tasks. The poor performance may suggest that the specific characteristics of the nephrology dataset, or the hyperparameters used, do not align well with the strengths of this gradient boosting framework. Further investigation into feature importance and model tuning would be necessary to understand this anomaly.
3. AdaBoost – Adaptive Boosting yielded a precision score of 0.89, demonstrating its capability to enhance weak classifiers by focusing on misclassified instances. This performance indicates that AdaBoost effectively leveraged the synthetic data augmentation and was able to generalize well across the test dataset.
4. Support Vector Machine – the SVM classifier, with a precision score of 0.93, performed robustly, confirming its strength in high-dimensional spaces, where it constructs optimal hyperplanes to segregate different risk classes. This suggests that the SVM is particularly well-suited to handle the complexity of the feature space derived from nephrology patient data.
5. Random Forest – the RF model achieved the highest precision score of 0.98. This ensemble method's exceptional performance underscores its ability to manage the variability within the dataset by combining the predictions of multiple decision trees, thereby enhancing overall predictive accuracy.
6. Decision Tree – the single DT model also performed well, with a precision score of 0.96. Despite its simplicity compared to ensemble methods, the decision tree's interpretability and effectiveness in handling this particular dataset are evident from its high score.
7. K-Nearest Neighbors – KNN classifier attained a precision score of 0.88. While KNN is often sensitive to the local structure of the data and can be affected by noise, its performance in this scenario indicates a reasonable degree of success in classifying risk levels among nephrology patients.

In summary, the Random Forest model emerged as the most effective classifier in terms of precision, followed closely by the Decision Tree and MLP classifiers. The unusually low precision score of LightGBM warrants further exploration to identify potential causes and corrective measures. These findings will inform future work in refining model selection and optimization for nephrology risk classification tasks.

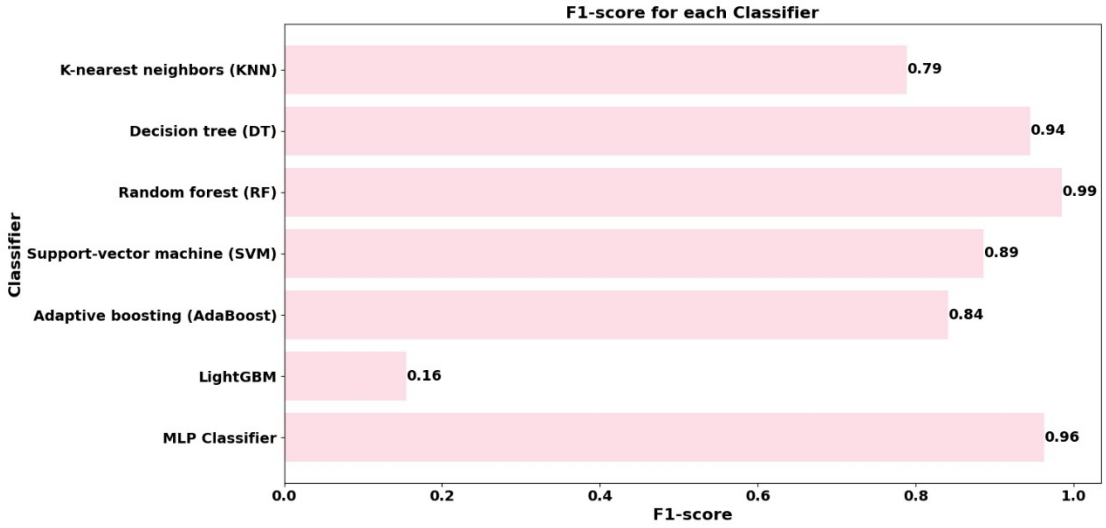

Figure 4: F1-score results for tested models.

The F1-score evaluation of different classifiers (see Fig. 4), as presented in the nephrology risk classification study, provides deeper insights into the balance between precision and recall across various models. Here, we discuss the performance of each classifier in terms of its F1-score, which is particularly useful in assessing the model's effectiveness in dealing with the complexities of an imbalanced dataset:

1. MLP Classifier – F1-score of 0.96 confirms strong capability in managing complex, non-linear classification, balancing precision and recall effectively
2. LightGBM – low F1-score of 0.16, suggesting significant struggles with both precision and recall, making it less suitable for this task.
3. Adaptive Boosting – F1-score of 0.84 reflects good performance, particularly in enhancing classification of difficult instances in imbalanced data.
4. Support Vector Machine – SVM scored 0.89 in F1, indicating strong performance with well-balanced precision and recall, effectively separating risk classes.
5. Random Forest – excelled with an F1-score of 0.99, showcasing superior performance by combining multiple decision trees to minimize errors.
6. Decision Tree – DT achieved a high F1-score of 0.94, demonstrating effective classification by accurately partitioning the feature space.
7. K-Nearest Neighbors – F1-score of 0.79 indicates moderate performance, reflecting challenges in balancing precision and recall, particularly with imbalanced data.

In conclusion, the Random Forest classifier once again proved to be the most effective model in terms of the F1-score, closely followed by the MLP and Decision Tree classifiers. The consistently low performance of LightGBM, as reflected in both precision and F1-score, warrants further exploration. Overall, these results provide clear guidance on the most appropriate machine learning models for nephrology risk classification tasks, with a particular emphasis on the effectiveness of ensemble methods and neural network-based approaches.

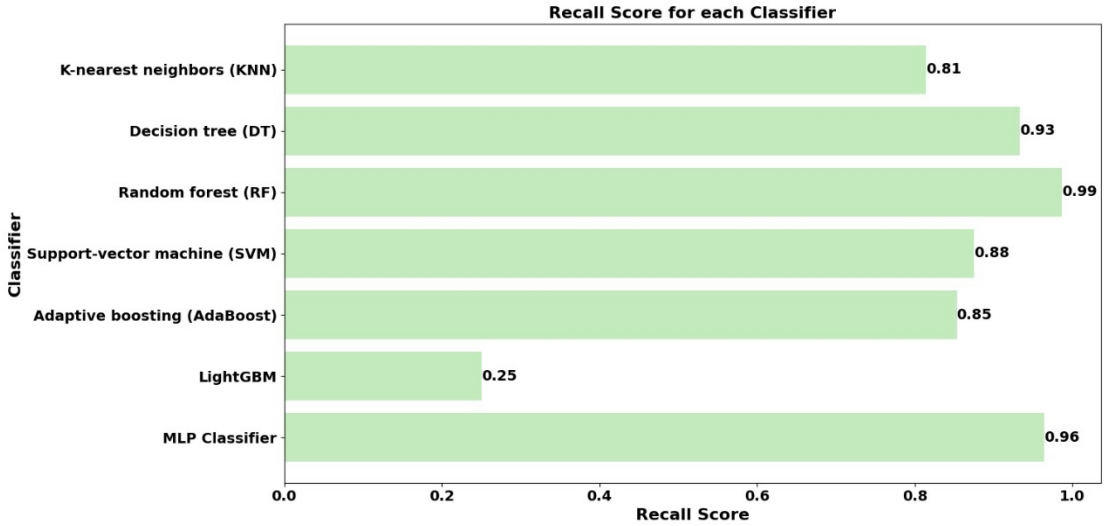

Figure 5: Recall results for tested models.

The recall scores (see Fig. 5) across various classifiers in the context of nephrology risk classification provide key insights into each model's ability to identify relevant cases within the dataset. Below is a summary of the performance for each classifier:

1. MLP Classifier – the MLP achieved a high recall score of 0.96, confirming its strong ability to identify nearly all relevant cases, making it one of the most effective models in this study.
2. LightGBM – the LightGBM model exhibited a low recall score of 0.25, indicating significant issues with sensitivity and a tendency to miss a substantial number of relevant cases.
3. Adaptive Boosting – AdaBoost scored 0.85 in recall, suggesting good sensitivity, as it effectively identified most of the relevant instances in the dataset.
4. Support Vector Machine – the SVM achieved a recall score of 0.88, showing strong sensitivity and a solid ability to correctly identify relevant cases, with few false negatives.
5. Random Forest – with an exceptional recall score of 0.99, the Random Forest model nearly perfectly identified relevant instances, making it highly reliable in this classification task.
6. Decision Tree – the DT model recorded a high recall score of 0.93, demonstrating its effectiveness in capturing most relevant cases and minimizing false negatives.
7. K-Nearest Neighbors – KNN classifier achieved a recall score of 0.81, indicating decent sensitivity but missing some relevant instances, reflecting moderate recall performance.

In summary, the Random Forest model again demonstrated the highest reliability with its near-perfect recall score, followed closely by the MLP and Decision Tree classifiers. The low recall score of LightGBM, similar to its performance in other metrics, suggests that it struggled significantly with this specific task.

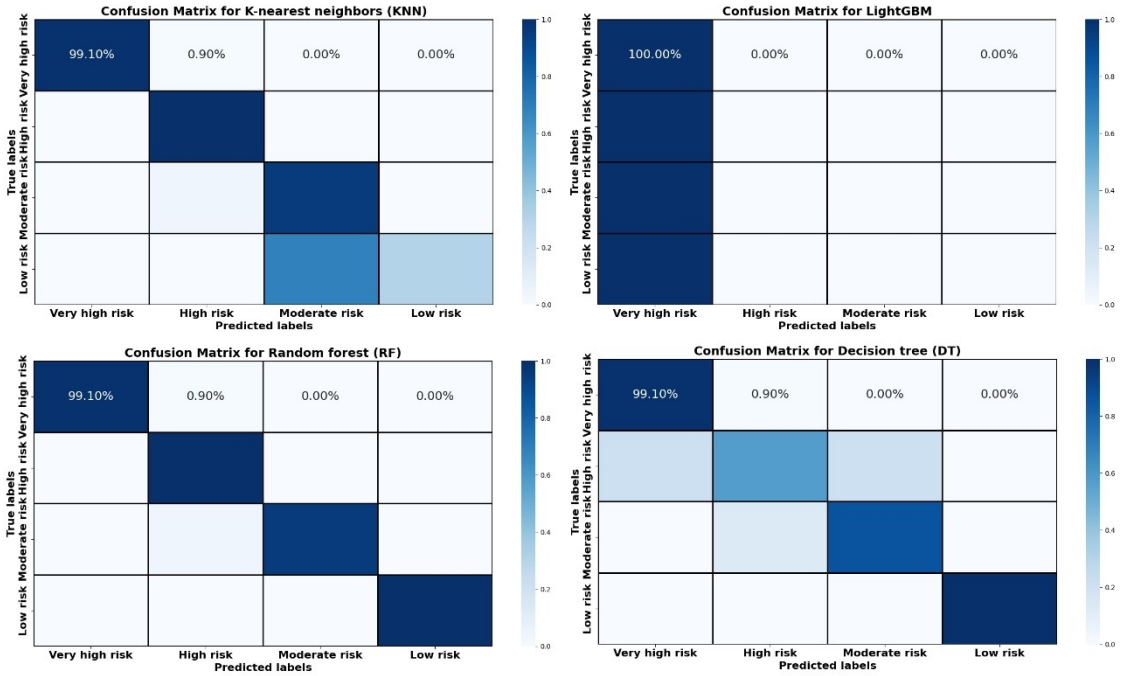

Figure 6: Example results of the confusion matrix for models.

The confusion matrices (see Fig. 6) presented in the provided figure summarize the performance of four different machine learning algorithms – KNN, LightGBM, RF, and DT.

The results presented in Figure 6 clearly indicate that the RF algorithm in the classification model achieved the best results in the course of the study. On the other hand, LightGBM, similarly to the graphics presented in Figures 3-5, performed the worst, which is confirmed by the result on the confusion matrix. In the case of the other two algorithms – KNN and DT, their efficiency in the task of risk classification of patients with IMN is confirmed to be decent, but not as good as for RF.

6.Conclusions

The comparative analysis of various artificial intelligence and machine learning algorithms for the classification of idiopathic membranous nephropathy risk demonstrates the potential of these technologies in enhancing traditional nephrological practices. The study evaluated seven different AI models, including K-Nearest Neighbors, Decision Trees, Random Forests, Support Vector Machines, Adaptive Boosting, LightGBM, and Multi-layer Perceptron Classifier.

The results indicate that ensemble methods such as Random Forests outperform other models, achieving a precision score of 0.98, demonstrating their robustness in handling the variability within nephrological data. The MLP Classifier and Decision Tree also showed high precision scores (0.96), suggesting their capability in capturing complex relationships within the clinical dataset. In contrast, LightGBM's performance was unexpectedly low, with a precision score of 0.11, indicating that further tuning and investigation into model parameters are required for this specific application.

The study highlights that AI models, particularly ensemble methods, can significantly improve the accuracy of risk classification in nephrology, potentially leading to better patient

outcomes. However, the results also underscore the importance of selecting appropriate models and fine-tuning them to the specific characteristics of medical datasets. Future research should focus on optimizing these models and exploring their integration into clinical decision support systems to assist nephrologists in managing kidney diseases more effectively.

Overall, this work provides a foundation for the continued exploration and application of AI in nephrology, emphasizing the need for further validation and refinement to achieve reliable and clinically applicable tools.

References

[1] Kovesdy, C. P. (2022). Epidemiology of chronic kidney disease: an update 2022, *Kidney International Supplements*, Vol. 12, No. 1, 7–11

[2] Niel, O.; Bastard, P. (2019). Artificial intelligence in nephrology: core concepts, clinical applications, and perspectives, *American Journal of Kidney Diseases*, Vol. 74, No. 6, 803–810

[3] Himmelfarb, J.; Vanholder, R.; Mehrotra, R.; Tonelli, M. (2020). The current and future landscape of dialysis, *Nature Reviews Nephrology*, Vol. 16, No. 10, 573–585

[4] Copur, S.; Tanriover, C.; Yavuz, F.; Soler, M. J.; Ortiz, A.; Covic, A.; Kanbay, M. (2023). Novel strategies in nephrology: what to expect from the future?, *Clinical Kidney Journal*, Vol. 16, No. 2, 230–244

[5] Pawuś, D.; Porażko, T.; Paszkiel, S. (2024). Automation and decision support in the area of nephrology using numerical algorithms, artificial intelligence and expert approachreview of the current state of knowledge, *IEEE Access*

[6] Rosner, M. H.; Berns, J. S. (2018). Transforming nephrology, *Clinical Journal of the American Society of Nephrology: CJASN*, Vol. 13, No. 2, 331

[7] Galuzio, P. P.; Cherif, A. (2022). Recent Advances and Future Perspectives in the Use of Machine Learning and Mathematical Models in Nephrology, *Advances in Chronic Kidney Disease*, Vol. 29, No. 5, 472–479

[8] Narayan, S.; others. (2022). A machine learning model for acute kidney injury prediction with novel kidney biomarkers, *2022 Second International Conference on Next Generation Intelligent Systems (ICNGIS)*, IEEE, 1–5

[9] Anand, S.; Verma, A. (2024). Artificial intelligence in nephrology, *Artificial Intelligence in Clinical Practice*, Elsevier, 201–209

[10] Bajaj, T.; Koyner, J. L. (2023). Cautious optimism: Artificial intelligence and acute kidney injury, *Clinical Journal of the American Society of Nephrology*, Vol. 18, No. 5, 668– 670

[11] Rank, N.; Pfahringer, B.; Kempfert, J.; Stamm, C.; Kühne, T.; Schoenrath, F.; Falk, V.; Eickhoff, C.; Meyer, A. (2020). Deep-learning-based real-time prediction of acute kidney injury outperforms human predictive performance, *NPJ Digital Medicine*, Vol. 3, No. 1, 139

[12] Chaudhari, A.; Sarode, V.; Udtewar, S.; Moharkar, L.; Patil, L.; Barreto, F. (2023). A Review of Artificial Intelligence for Predictive Healthcare Analytics and Healthcare IoT Applications, *International Conference on Intelligent Computing and Networking*, Springer, 555–562

[13] Zhang, H.; Preddie, D.; Krackov, W.; Sor, M.; Waguespack, P.; Kuang, Z.; Ye, X.; Kotanko, P. (2022). Deep learning to classify arteriovenous access aneurysms in hemodialysis patients, *Clinical Kidney Journal*, Vol. 15, No. 4, 829–830

[14] Chan, L.; Vaid, A.; Nadkarni, G. N. (2020). Applications of machine learning methods in kidney disease: hope or hype?, *Current Opinion in Nephrology and Hypertension*, Vol. 29, No. 3, 319–326

[15] Ho, C. W.-L.; Caals, K. (2021). A call for an ethics and governance action plan to harness the power of artificial intelligence and digitalization in nephrology, *Seminars in Nephrology* (Vol. 41), Elsevier, 282–293

[16] Miao, J.; Thongprayoon, C.; Suppadungsuk, S.; Garcia Valencia, O. A.; Qureshi, F.; Cheungpasitporn, W. (2023). Ethical Dilemmas in Using AI for Academic Writing and an Example Framework for Peer Review in Nephrology Academia: A Narrative Review, *Clinics and Practice*, Vol. 14, No. 1, 89–105

[17] Wei, W.; Cai, Z.; Chen, L.; Yuan, W.; Fan, Y.; Rong, S. (2023). Short-term prognostic models for severe acute kidney injury patients receiving prolonged intermittent renal replacement therapy based on machine learning, *BMC Medical Informatics and Decision Making*, Vol. 23, No. 1, 133

[18] Kusumadewi, S.; Rosita, L.; Wahyuni, E. G. (2023). Fuzzy linear regression based on a hybrid of fuzzy C-means and the fuzzy inference system for predicting serum iron levels in patients with chronic kidney disease, *Expert Systems with Applications*, Vol. 227, 120314

[19] Nakamura, K.; Uchino, E.; Sato, N.; Araki, A.; Terayama, K.; Kojima, R.; Murashita, K.; Itoh, K.; Mikami, T.; Tamada, Y.; others. (2023). Individual health-disease phase diagrams for disease prevention based on machine learning, *Journal of Biomedical Informatics*, Vol. 144, 104448

[20] Akter, S.; Ahmed, M.; Al Imran, A.; Habib, A.; Haque, R. U.; Rahman, M. S.; Hasan, M. R.; Mahjabeen, S. (2023). CKD. Net: A novel deep learning hybrid model for effective, real-time, automated screening tool towards prediction of multi stages of CKD along with eGFR and creatinine, *Expert Systems with Applications*, Vol. 223, 119851

[21] Kim, D.-W.; Ahn, H.-G.; Kim, J.; Yoon, C.-S.; Kim, J.-H.; Yang, S. (2021). Advanced kidney volume measurement method using ultrasonography with artificial intelligencebased hybrid learning in children, *Sensors*, Vol. 21, No. 20, 6846

[22] Singh, N. P.; Bapi, R. S.; Vinod, P. (2018). Machine learning models to predict the progression from early to late stages of papillary renal cell carcinoma, *Computers in Biology and Medicine*, Vol. 100, 92–99

[23] Karar, M. E.; El-Fishawy, N.; Radad, M. (2023). Automated classification of urine biomarkers to diagnose pancreatic cancer using 1-D convolutional neural networks, *Journal of Biological Engineering*, Vol. 17, No. 1, 28

[24] Chan, L.; Nadkarni, G. N.; Fleming, F.; McCullough, J. R.; Connolly, P.; Mosoyan, G.; El Salem, F.; Kattan, M. W.; Vassalotti, J. A.; Murphy, B.; others. (2021). Derivation and validation of a machine learning risk score using biomarker and electronic patient data to predict progression of diabetic kidney disease, *Diabetologia*, Vol. 64, 1504–1515

[25] Hu, C.; Tan, Q.; Zhang, Q.; Li, Y.; Wang, F.; Zou, X.; Peng, Z. (2022). Application of interpretable machine learning for early prediction of prognosis in acute kidney injury, *Computational and Structural Biotechnology Journal*, Vol. 20, 2861–2870

[26] Li, C.; Yao, Z.; Zhu, M.; Lu, B.; Xu, H. (2018). Biopsy-free prediction of pathologic type of primary nephrotic syndrome using a machine learning algorithm, *Kidney and Blood Pressure Research*, Vol. 42, No. 6, 1045–1052

[27] Pan, Y.-B.; Ye, H.-M.; Jiang, Z.-H.; Chen, D.-J.; Teng, Y.; Guan, C.-A. (2022). Identification and validation of two hub genes involved in membranous nephropathy based on machine learning

[28] Duo, L.; Chen, L.; Zuo, Y.; Guo, J.; He, M.; Zhao, H.; Kang, Y.; Tang, W. (2023). Machine learning model to estimate probability of remission in patients with idiopathic membranous nephropathy, *International Immunopharmacology*, Vol. 125, 111126

[29] Destere, A.; Teisseyre, M.; Merino, D.; Cremoni, M.; Gérard, A. O.; Crepin, T.; JourdeChiche, N.; Graça, D.; Zorzi, K.; Fernandez, C.; others. (2023). Optimization of rituximab

Therapy in Adult patients with PLA2R1-associated Membranous Nephropathy with Artificial Intelligence, *Kidney International Reports*

[30] Hao, F.; Liu, X.; Li, M.; Han, W. (2023). Accurate Kidney Pathological Image Classification Method Based on Deep Learning and Multi-Modal Fusion Method with Application to Membranous Nephropathy, *Life*, Vol. 13, No. 2, 399

[31] Gao, J.; Wang, S.; Xu, L.; Wang, J.; Guo, J.; Wang, H.; Sun, J. (2023). Computer-aided diagnosis of primary membranous nephropathy using expert system, *BioMedical Engineering OnLine*, Vol. 22, No. 1, 1–24

[32] Uleri, A.; Baboudjian, M.; Gallioli, A.; Territo, A.; Gaya, J. M.; Sanz, I.; Robalino, J.; Casadevall, M.; Diana, P.; Verri, P.; others. (2023). A new machine-learning model to predict long-term renal function impairment after minimally invasive partial nephrectomy: the Fundació Puigvert predictive model, *World Journal of Urology*, Vol. 41, No. 11, 2985–2990

[33] Kotanko, P.; Zhang, H.; Wang, Y. (2023). Artificial Intelligence and Machine Learning in Dialysis: Ready for Prime Time?, *Clinical Journal of the American Society of Nephrology*, 10–2215

[34] Luo, W.; Phung, D.; Tran, T.; Gupta, S.; Rana, S.; Karmakar, C.; Shilton, A.; Yearwood, J.; Dimitrova, N.; Ho, T. B.; others. (2016). Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view, *Journal of Medical Internet Research*, Vol. 18, No. 12, e323

[35] Van Smeden, M.; Reitsma, J. B.; Riley, R. D.; Collins, G. S.; Moons, K. G. (2021). Clinical prediction models: diagnosis versus prognosis, *Journal of Clinical Epidemiology*, Vol. 132, 142–145

[36] Collins, G. S.; Reitsma, J. B.; Altman, D. G.; Moons, K. G. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) the TRIPOD statement, *Circulation*, Vol. 131, No. 2, 211–219

[37] Riley, R. D.; Ensor, J.; Snell, K. I.; Harrell, F. E.; Martin, G. P.; Reitsma, J. B.; Moons, K. G.; Collins, G.; Van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model, *Bmj*, Vol. 368

[38] Christodoulou, E.; van Smeden, M.; Edlinger, M.; Timmerman, D.; Wanitschek, M.; Steyerberg, E. W.; Van Calster, B. (2021). Adaptive sample size determination for the development of clinical prediction models, *Diagnostic and Prognostic Research*, Vol. 5, No. 1, 1–12

[39] Riley, R. D.; Debray, T. P.; Collins, G. S.; Archer, L.; Ensor, J.; van Smeden, M.; Snell, K. I. (2021). Minimum sample size for external validation of a clinical prediction model with a binary outcome, *Statistics in Medicine*, Vol. 40, No. 19, 4230–4251

[40] Archer, L.; Snell, K. I.; Ensor, J.; Hudda, M. T.; Collins, G. S.; Riley, R. D. (2021). Minimum sample size for external validation of a clinical prediction model with a continuous outcome, *Statistics in Medicine*, Vol. 40, No. 1, 133–146

[41] Rovin, B. H.; Adler, S. G.; Barratt, J.; Bridoux, F.; Burdge, K. A.; Chan, T. M.; Cook, H. T.; Fervenza, F. C.; Gibson, K. L.; Glassock, R. J.; others. (2021). KDIGO 2021 clinical practice guideline for the management of glomerular diseases, *Kidney International*, Vol. 100, No. 4, S1–S276