Decision Tree for Predicting the Mortality in Hemodialysis Patient with Diabetes

[1]Noper Ardi, [2]Ahmadi Irmansyah Lubis, [3]Isnayanti
[1, 2]Politeknik Negeri Batam, Indonesia, [3]Universitas Terbuka, Indonesia

[1]noperardi@polibatam.ac.id, [2]ahmadi@polibatam.ac.id, [3]isnayanti30@gmail.com,

ABSTRAK

Hemodialysis patients with diabetes face a significantly higher risk of mortality compared to those without diabetes. Accurate prediction of mortality in this patient population is crucial for guiding clinical decision-making, improving patient outcomes, and optimizing resource allocation. Hemodialysis is a procedure for cleaning the blood from the waste products of the body's metabolism. this is one of modality to treat end stage kidney disease. Diabetes mellitus is a significant contributor to the global burden of chronic kidney disease (CKD), and patients with diabetes undergoing hemodialysis are at a higher risk of mortality compared to those without diabetes. Identifying factors that influence mortality risk in this population can aid in clinical decision-making and improve patient outcomes. Dialysis is performed on patients with kidney failure, both acute kidney failure and chronic kidney failure. This study is aimed to predict the mortality risk of hemodialysis patients with diabetes. The Taiwanese hemodialysis center enrolled a total of 665 hemodialysis patients. The prediction is based on Decision Tree. Compared with K-Nearest Neighbor, linear discriminant, Logistic Regression, and Ensemble, Decission Tree performed better. As for related medical variables like parathyroid surgery, urea reduction ratio, etc., they play a much smaller role in mortality risk factors than diabetes and cardiovascular disease.

Key Words: Decision Tree, Prediction, Hemodialysis, Chronic kidney disease, Diabetes

INTRODUCTION

Chronic kidney disease (CKD) is a complex and prevalent condition that affects millions of individuals worldwide. Among CKD patients, those with diabetes mellitus face a considerably higher risk of mortality. Hemodialysis is a commonly utilized renal replacement therapy for patients with CKD, but identifying the factors that contribute to mortality in hemodialysis patients with diabetes remains a challenge. Accurate prediction of mortality risk can significantly impact clinical decision-making, allowing for timely interventions and improved patient outcomes(X. Liu et al. 2021).

One promising approach for mortality prediction in healthcare is the use of decision tree algorithms. Decision trees are powerful predictive models that use a hierarchical structure of binary decisions to classify data based on input variables(Borisagar, Barad, and Raval 2017). They offer interpretability, simplicity, and the ability to handle both categorical and continuous data. In the context of predicting mortality in hemodialysis patients with diabetes, decision trees can effectively identify the key risk factors and facilitate clinical decision-making(Yeh, Wu, and Tsao 2011).

This paper aims to explore the application of decision tree algorithms for predicting mortality in hemodialysis patients with diabetes. By examining relevant literature and studies, we will analyze the strengths and limitations of decision trees in this specific healthcare context. Additionally, we will evaluate the performance of decision tree models against other predictive

models, such as logistic regression or neural networks, to determine their efficacy in mortality prediction(Liu et al. 2018).

The utilization of decision trees for mortality prediction in hemodialysis patients with diabetes offers several advantages. Firstly, decision trees provide a transparent and interpretable framework, allowing clinicians to understand the decision-making process and identify critical factors contributing to mortality risk(Huda and Ardi 2021). This interpretability is crucial in medical settings, where trust and understanding of the model's predictions are paramount. Furthermore, decision trees can handle missing values and outliers, making them robust in real-world clinical scenarios where data quality may vary(Ardi, Setiawan, and Bharata Adji 2019).

To achieve the objectives of this paper, we will review and analyze relevant studies that have employed decision tree algorithms for mortality prediction in hemodialysis patients with diabetes. We will explore the different input variables considered in these models, such as demographic information, laboratory values, comorbidities, and dialysis-related parameters. Moreover, we will assess the performance metrics, including accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC), to evaluate the predictive capabilities of decision tree models(Y. S. Liu et al. 2021).

In conclusion, predicting mortality in hemodialysis patients with diabetes is a critical task that can aid in clinical decision-making and improve patient care. Decision tree algorithms offer a promising approach due to their interpretability and ability to handle complex healthcare datasets. By examining the strengths and limitations of decision trees in this context, this paper aims to contribute to the existing body of knowledge and inform future research and clinical practice in mortality prediction for hemodialysis patients with diabetes.

METHODOLOGY

Kidney Failure

Kidney failure, also known as renal failure, is a condition in which the kidneys are no longer able to function properly. The kidneys are a pair of organs located in the abdomen that play a vital role in filtering waste products and excess fluid from the blood, regulating electrolyte balance, and producing hormones that help regulate blood pressure and stimulate red blood cell production(Mirinejad and Inanc 2015).

There are two main types of kidney failure: acute kidney injury (AKI) and chronic kidney disease (CKD). AKI is a sudden and rapid loss of kidney function that typically occurs over a few hours or days. It can be caused by a variety of factors, such as dehydration, low blood volume, infections, medications, toxins, and kidney damage from a traumatic injury or surgery. Symptoms of AKI may include decreased urine output, swelling in the legs or feet, fatigue, confusion, nausea, and shortness of breath(X. Liu et al. 2021). Treatment for AKI may include managing the underlying cause, such as stopping a medication or treating an infection, and providing supportive care, such as dialysis, to help the kidneys recover(Panwong and Iam-On 2016).

CKD is a progressive loss of kidney function that occurs over a period of months or years. It is typically caused by underlying health conditions, such as diabetes, high blood pressure, and autoimmune disorders, that damage the kidneys over time. Symptoms of CKD may not appear until the disease is in advanced stages, and may include fatigue, itching, swelling in the legs or feet, loss of appetite, and changes in urine output or color(Drüeke et al. 2006). Treatment for CKD may include managing the underlying health conditions, such as controlling blood sugar and blood pressure, and lifestyle changes, such as a healthy diet and regular exercise. In advanced stages of CKD, dialysis or kidney transplantation may be necessary(Patil 2017).

Kidney failure can lead to a number of serious complications, such as anemia, bone disease, high blood pressure, heart disease, and nerve damage. It can also affect the body's ability to regulate electrolyte balance, which can cause imbalances in potassium, sodium, and calcium levels. In severe cases, kidney failure can be life-threatening and may require hospitalization and intensive treatment(Garcia-Montemayor et al. 2021).

Prevention of kidney failure involves managing underlying health conditions that can damage the kidneys, such as diabetes, high blood pressure, and autoimmune disorders. Adopting a healthy lifestyle, such as a balanced diet and regular exercise, can also help prevent kidney damage(Somji, Ruggajo, and Moledina 2020). It is important to avoid medications and substances that can be toxic to the kidneys, such as certain pain medications, and to have regular check-ups with a healthcare provider to monitor kidney function(Escandell-Montero et al. 2014).

Hemodialisys

Hemodialysis is a medical procedure that is used to remove waste products and excess fluid from the blood when the kidneys are no longer able to do so. It is a form of renal replacement therapy and is one of the most common treatments for end-stage renal disease (ESRD), a condition in which the kidneys have failed and can no longer function properly(Sung et al. 2014).

During hemodialysis, blood is pumped out of the patient's body through a vascular access point, such as an arteriovenous fistula or graft, or a central venous catheter. The blood is then filtered through a dialyzer, a machine that contains a semipermeable membrane that allows small molecules such as waste products and excess fluids to pass through, while retaining larger molecules such as proteins and blood cells(Escandell-Montero et al. 2014).

The dialyzer is connected to the patient's bloodstream through two needles or catheters, one to draw blood out of the body and another to return the filtered blood to the patient. The blood flows through the dialyzer in one direction, while a dialysis solution, also called dialysate, flows in the opposite direction on the other side of the membrane. The dialysis solution contains electrolytes, such as sodium and potassium, and is used to maintain the proper balance of these substances in the patient's blood(Lacson 2008).

During the hemodialysis session, the patient's blood is continuously filtered through the dialyzer for several hours, usually three times a week, with each session lasting between 3 and 5 hours, depending on the patient's condition. The duration of the session can vary depending on the amount of waste producs and excess fluids that need to be removed from the patient's blood(Kim et al. 2021).

Hemodialysis is a complex procedure that requires specialized equipment and trained healthcare professionals, such as nurses and technicians. The process requires careful monitoring of the patient's vital signs, such as blood pressure, heart rate, and oxygen levels, to ensure that the procedure is safe and effective(M. Kumar 2016).

In addition to removing waste products and excess fluids from the blood, hemodialysis can also be used to treat certain medical conditions, such as acute kidney injury, drug overdose, and certain types of poisoning. However, hemodialysis is not a cure for kidney disease and does not restore the kidneys to normal function. It is a lifelong treatment for ESRD and requires ongoing management and monitoring by a healthcare team(S.Ramya 2016).

Decision Tree Algorithm

The decision tree algorithm is a machine learning technique used for classification and regression tasks. It builds a tree-like model of decisions and their possible consequences by recursively splitting the dataset based on different features or attributes(Potharaju and Sreedevi 2016)(Ardi and Isnayanti 2020).

The decision tree algorithm works, with the following steps:

1. Data Preparation: Start with a labeled dataset, where each instance has a set of input features and corresponding target labels or outcomes.
2. Attribute Selection: Choose the attribute that best separates the data based on a certain criterion. Common criteria include Gini impurity and information gain. The attribute with the highest impurity reduction or information gain is selected for splitting.
3. Splitting: Split the dataset based on the chosen attribute into subsets or branches. Each branch represents a unique value or category of the attribute. Instances are allocated to the appropriate branch based on their attribute values.

4. Recursion: Repeat the splitting process recursively on each branch, considering the remaining attributes and subset of instances at each node. This process continues until a stopping condition is met, such as reaching a predefined depth limit, having a minimum number of instances in a node, or achieving a pure subset (all instances have the same label).

5. Leaf Node Assignment: Assign a class or outcome label to each leaf node. For classification tasks, the label can be the majority class in the corresponding subset. For regression tasks, the label can be the mean or median value of the target variable in the subset.

6. Pruning (optional): After building the initial decision tree, pruning can be applied to reduce overfitting. Pruning involves removing unnecessary branches or nodes from the tree based on performance metrics, such as validation accuracy or error rates.

7. Prediction: To make predictions for new instances, follow the decision path from the root node to a leaf node based on the attribute values of the instance. The predicted outcome is the label assigned to the leaf node.

8. The decision tree algorithm offers several advantages. It is easily interpretable and provides insights into the decision-making process. Decision trees can handle both categorical and continuous features, as well as missing data. They are also relatively fast in terms of training and prediction.

However, decision trees are prone to overfitting and can be sensitive to small changes in the data. They may not always generalize well to unseen instances. To address these limitations, techniques like pruning, setting depth limits, or using ensemble methods like random forests or gradient boosting can be employed. Overall, the decision tree algorithm is widely used in various domains due to its simplicity, nterpretability, and ability to handle different types of data(Deshwal, Sangwan, and Kumar 2020).

RESULT AND DISCUSSION

1. Prediction Results of Dataset A

Based on the results of Decision Tree (DT) testing, it was observed that Dataset A had a correct rate (acc) training of 0.833 for a period of 334.51 msec. In comparison with KNN, Logistic Regression, and linear discriminant (LD), the Decision Tress execution time is in top three just below KNN. A linear discriminant and ensemble logic algorithm achieved a short training time of 318.69 msec, and a long training time of 695.13 msec for dataset A respectively. Dataset A is listed in Table 1 along with the complete training comparison results.
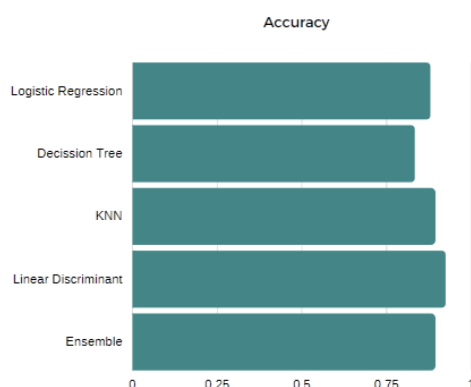

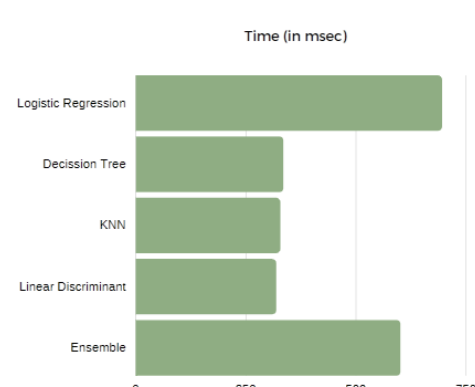Fig 1. Training comparison of Dataset A -Accuracy   Fig 2. Training comparison of Dataset A – time

An analysis of the data using Decision Tree yielded a correct rate of 0.833. The Decision Three Has average performance than Linear Discriminant, Logistic Regression, Ensemble (En), and KNN. In Table 1, we present the detail predictions based on dataset A. SVM presents the highest accuracy in predicting mortality for Dataset A, 0.939, followed by Linear Discriminant 0.924, Ensemble and KKN 0.894, Logistic Regression 0.879, and Decision Tree 0.833.

Table 1. Prediction Result Comparison of Dataset A

Metrics	LR	DT	KNN	LD	En
Correct rate	0.960	0.924	0.893	0.955	0.894
Error rate	0.040	0.076	0.107	0.045	0.106
Sensitivity (TPR)	0.569	0.429	0.514	0.571	0
Specificity (TNR)	1	0.983	0.974	1	1
Positive Predictive Value	1	0.750	1	1	NaN
Negative Predictive Value	0.952	0.936	1	0.952	0.894
PositiveLikelihood	NaN	25.286	NaN	NaN	NaN
NegativeLikelihood	0	0.581	0	0.429	1
Prevalence	0.106	0.106	0.106	0.106	0.106
F1	0.725	0.546	0.679	0.727	NaN
Bookmaker Informedness (BM)	0.569	0.412	0.488	0.571	0
Markedness(MK)	0.952	0.686	1	0.952	NaN

Table 1 shows that the highest correct rate of prediction for Dataset A is 0,960 for Logistic Regression, followed by 0.955 for Linear Discrimination, 0.924 for Decision Tree, 0.894 for Ensemble, and 0.893 for KKN.

2. Prediction Results of Dataset B

Of the total 354 patients Dataset B, Decision Tree successfully predicted with precision reaching 0.9270 and up to 0.944 respectively for training and test. Decision Tree has the best performance compared to other methods.
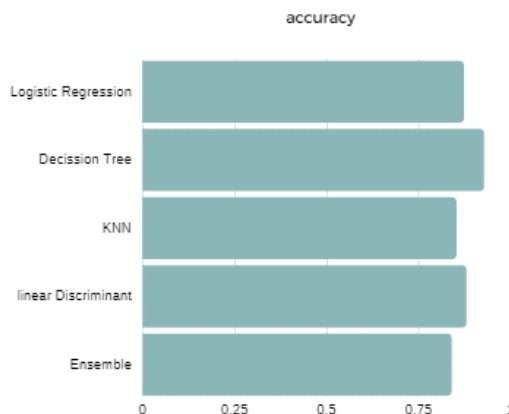

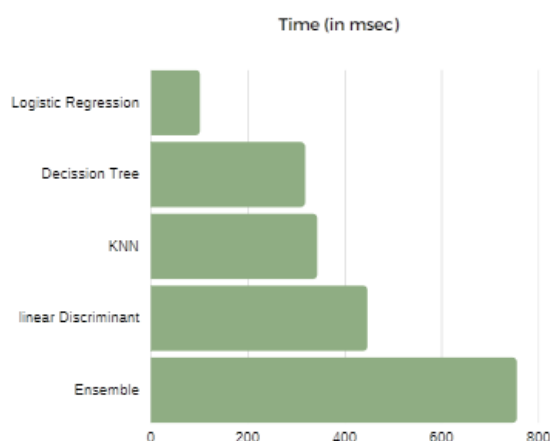Fig 3. Training comparison of Dataset B - Accuracy

Fig 4. Training comparison of Dataset B - time

Using Dataset B, as shown in figure 3 and figure 4, it is evident that the Decision Tree classifier shows the highest accuracy in predicting the mortality in Hemodialysis patients with diabetes, i.e.at the rate of 0.927, followed by Linear Discriminant of 0.879, Logistic Regression of 0.873, KKN of 0.853, and the last one is Ensemble of 0.839.

Table 2. Prediction Result Comparison of Dataset B

Metrics	LR	DT	KNN	LD	En
Correct rate	0.893	0.944	0.880	0.901	0.893
Error rate	0.107	0.057	0.120	0.099	0.107
Sensitivity (TPR)	0.368	0.825	0.544	0.544	0.368
Specificity (TNR)	0.993	0.966	0.970	0.970	0.993
Positive Predictive Value	0.913	0.825	0.775	0.775	0.913
Negative Predictive Value	0.891	0.966	0.917	0.917	0.891
PositiveLikelihood	54.710	24.490	NaN	17.947	54.710
NegativeLikelihood	0.636	0.182	0	0.470	0.636
Prevalence	0.161	0.161	0.161	0.161	0.161
F1	0.693	0.825	0.693	0.693	0.693
Bookmaker Informedness (BM)	0.362	0.791	0.514	0.514	0.362
Markedness(MK)	0.804	0.791	0.692	0.692	0.804

Using Dataset B, as shown in Table 2, it is evident that the Decision Tree classifier shows the highest correct rate in predicting the mortality of hemodialysis patients with diabetes, i.e. at the rate of 0.944, followed by Linear Discriminant of 0.901, Logistic Regression and Ensemble of 0.893, and the last one is KKN of 0.880.

3. Prediction Results of Dataset C

After training on 170 hemodialysis patients in Dataset C, it was found that Decision Tree yielded accuracy of 0.771 achieved for 819.06 msec. Logistic Regression gives accuracy 0.818. KNN and Ensemble each provide accuracy of 0.765, and 0.641, respectively. Figure 5 and figure 6 shows the training result of Dataset C, whereas Table 3 shows the performances comparison results of Dataset C.
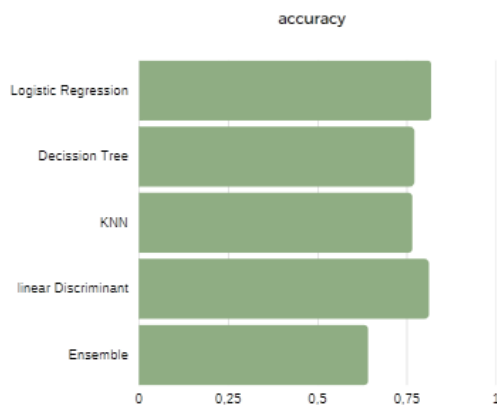

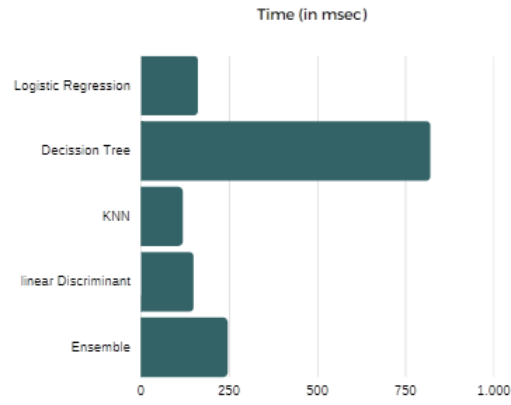Fig 5. Training comparison of Dataset C – Accurac   Fig 6. Training comparison of Dataset C - Time

Table 3. Prediction Result Comparison of Dataset C

Metrics	LR	DT	KNN	LD	En
Correct rate	0.882	0.859	0.841	0.841	0.641
Error rate	0.118	0.141	0.159	0.159	0.359
Sensitivity (TPR)	0.803	0.689	0.803	0.721	0
Specificity (TNR)	0.927	0.954	0.927	0.908	1
Positive Predictive Value	0.860	0.894	0.860	0.815	NaN
Negative Predictive Value	0.894	0.846	0.894	0.853	0.641
Positive Likelihood	10.945	15.010	10.945	7.862	NaN
Negative Likelihood	0.212	0.3275	0.212	0.307	1
Prevalence	0.359	0.359	0.359	0.359	0.359
F1	0.831	0.778	0.831	0.765	NaN
Bookmaker Informedness (BM)	0.730	0.643	0.730	0.630	0
Markedness(MK)	0.753	0.739	0.753	0.668	NaN

4. Prediction Results of Dataset D

Figure 7 and figure 8 show Decision Tree performance and comparison methods for the D dataset, respectively. After the training and the test turns out Decision Tree is still superior to its

competitors. Decision Tree provides accuracy at 0.720. The required execution time is 391.33 msec, faster than other methods.
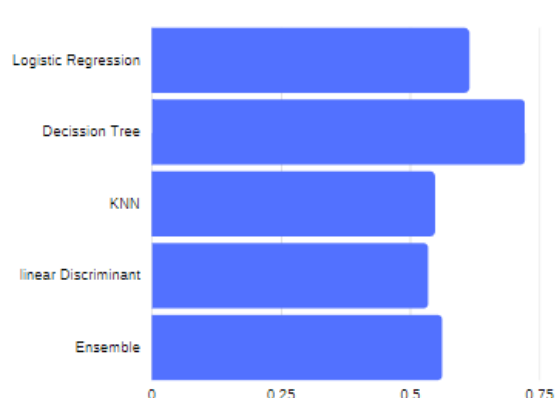

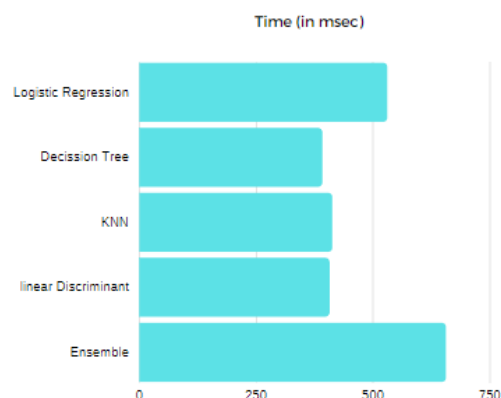Fig 7. Training comparison of Dataset D – Accuracy

Fig 8. Training comparison of Dataset D – Time

Table 4. Prediction Result Comparison of Dataset D

Metrics	LR	DT	KNN	LD	En
Correct rate	0.867	0.977	0.773	0.708	0.560
Error rate	0.133	0.023	0.227	0.293	0.440
Last Correct Rate	0.867	0.987	0.773	0.708	0.560
Last Error Rate	0.133	0.1330	0.227	0.293	0.440
Classified Rate	1	1	1	1	1
Sensitivity (TPR)	0.857	1	0.810	0.691	1
Specificity (TNR)	0.879	0.970	0.727	0.727	0
Positive Predictive Value	0.900	0.970	0.791	0.763	0.560
Negative Predictive Value	0.829	1	0.750	0.649	NaN
PositiveLikelihood	7.071	33	2.968	2.532	1
NegativeLikelihood	0.163	0	0.262	0.426	NaN
Prevalence	0.560	0.560	0.560	0.560	0.560
F1	0.878	0.985	0.800	0.725	NaN
Bookmaker Informedness (BM)	0.736	0.970	0.537	0.418	0
Markedness(MK)	0.729	0.970	0.541	0.412	NaN

Comparatively, to other algorithms, Decision Trees provide the highest training accuracy for Dataset D for predicting mortality in Hemodialisys patients with diabetes. A Logistic

Regression score of 0.613, Ensemble score of 0.560, KKN score of 0.547, and Linear Discriminant score of 0.533 followed by a Logistic Regression score of 0.613, Ensemble score of 0.560, KKN score of 0.547. A Decision Tree with a correct rate of 0.977 has the highest correct rate followed by Logistic Regression with 0.867, KKN in third with 0.7733 followed by Linear Discriminant with 0.7076, and Ensemble in fourth with 0.56.
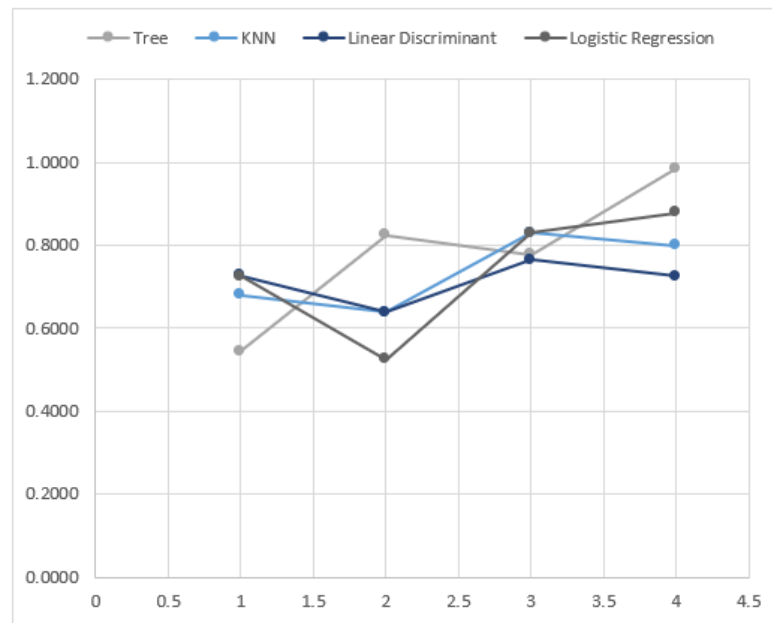

Fig 9. F1 Score comparison

CONCLUSION

In this paper, we explored the application of decision tree algorithms for predicting mortality in hemodialysis patients with diabetes. The use of decision trees in healthcare has gained significant attention due to their interpretability and ability to handle complex datasets. Our analysis revealed that decision trees offer a transparent and interpretable framework for mortality prediction, enabling clinicians to understand the factors influencing patient outcomes. The hierarchical structure of decision trees allows for the identification of important risk factors and their interactions, leading to actionable insights for clinical decision-making. Moreover, decision trees can handle both categorical and continuous variables, making them suitable for integrating diverse patient characteristics and biomarkers. However, it is important to acknowledge the limitations of decision trees. Overfitting remains a concern, as decision trees can capture noise or irrelevant patterns in the data. Pruning techniques and ensemble methods, such as random forests, can be employed to mitigate overfitting and enhance generalization. Additionally, the interpretability of decision trees may come at the expense of predictive power compared to more complex models like neural networks. Overall, decision trees provide a valuable tool in the realm of predictive medicine, offering insights into mortality risk assessment and contributing to personalized patient care for hemodialysis patients with diabetes.

REFERENCE

Ardi, Noper, and Isnayanti. 2020. "Structural Equation Modelling-Partial Least Square to Determine the Correlation of Factors Affecting Poverty in Indonesian Provinces." *IOP Conference Series: Materials Science and Engineering* 846(1):0–13. doi: 10.1088/1757-899X/846/1/012054.
Ardi, Noper, Noor Akhmad Setiawan, and Teguh Bharata Adji. 2019. "Analytical Incremental

Learning for Power Transformer Incipient Fault Diagnosis Based on Dissolved Gas Analysis." *Proceedings - 2019 5th International Conference on Science and Technology, ICST 2019* 3–6. doi: 10.1109/ICST47872.2019.9166441.

Borisagar, Nilesh, Dipa Barad, and Priyanka Raval. 2017. "Chronic Kidney Disease Prediction Using Back Propagation Neural Network Algorithm." *Advances in Intelligent Systems and Computing* 508:295–303. doi: 10.1007/978-981-10-2750-5_31.

Deshwal, Deepti, Pardeep Sangwan, and Divya Kumar. 2020. "A Language Identification System Using Hybrid Features and Back-Propagation Neural Network." *Applied Acoustics* 164:107289. doi: 10.1016/j.apacoust.2020.107289.

Drüeke, Tilman B., Francesco Locatelli, Naomi Clyne, Kai-Uwe Eckardt, Iain C. Macdougall, Dimitrios Tsakiris, Hans-Ulrich Burger, and Armin Scherhag. 2006. "Normalization of Hemoglobin Level in Patients with Chronic Kidney Disease and Anemia." *New England Journal of Medicine* 355(20):2071–84. doi: 10.1056/nejmoa062276.

Escandell-Montero, Pablo, Milena Chermisi, José M. Martínez-Martínez, Juan Gómez-Sanchis, Carlo Barbieri, Emilio Soria-Olivas, Flavio Mari, Joan Vila-Francés, Andrea Stopper, Emanuele Gatti, and José D. Martín-Guerrero. 2014. "Optimization of Anemia Treatment in Hemodialysis Patients via Reinforcement Learning." *Artificial Intelligence in Medicine* 62(1):47–60. doi: 10.1016/j.artmed.2014.07.004.

Garcia-Montemayor, Victoria, Alejandro Martin-Malo, Carlo Barbieri, Francesco Bellocchio, Sagrario Soriano, Victoria Pendon-Ruiz de Mier, Ignacio R. Molina, Pedro Aljama, and Mariano Rodriguez. 2021. "Predicting Mortality in Hemodialysis Patients Using Machine Learning Analysis." *Clinical Kidney Journal* 14(5):1388–95. doi: 10.1093/ckj/sfaa126.

Huda, Asrul, and Noper Ardi. 2021. "Predictive Analytic on Human Resource Department Data Based on Uncertain Numeric Features Classification." *International Journal of Interactive Mobile Technologies* 15(8):172–81. doi: 10.3991/ijim.v15i08.20907.

Kim, Hyung Woo, Seok Jae Heo, Jae Young Kim, Annie Kim, Chung Mo Nam, and Beom Seok Kim. 2021. "Dialysis Adequacy Predictions Using a Machine Learning Method." *Scientific Reports* 11(1). doi: 10.1038/s41598-021-94964-1.

Lacson, Ronilda. 2008. "Predicting Hemodialysis Mortality Utilizing Blood Pressure Trends." *AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium* 369–73.

Liu, Xiaobin, Xiran Zhang, Xiaoyi Guo, Yijie Ding, Weiwei Shan, Liang Wang, Wei Zhou, and Hua Shi. 2021. "A Self-Representation-Based Fuzzy SVM Model for Predicting Vascular Calcification of Hemodialysis Patients." *Computational and Mathematical Methods in Medicine* 2021. doi: 10.1155/2021/2464821.

Liu, Yi Shiuan, Chih Yu Yang, Ping Fang Chiu, Hui Chu Lin, Chung Chuan Lo, Alan Szu Han Lai, Chia Chu Chang, and Oscar Kuang Sheng Lee. 2021. "Machine Learning Analysis of Time-Dependent Features for Predicting Adverse Events during Hemodialysis Therapy: Model Development and Validation Study." *Journal of Medical Internet Research* 23(9). doi: 10.2196/27098.

Liu, Zhenyuan, Zaisheng Liu, Zhipeng Zhu, Yindong Shen, and Junwu Dong. 2018. "Simulated Annealing for a Multi-Level Nurse Rostering Problem in Hemodialysis Service." *Applied Soft Computing Journal* 64:148–60. doi: 10.1016/j.asoc.2017.12.005.

M. Kumar. 2016. "Prediction of Chronic Kidney Disease Using Random Forest Machine Learning Algorithm." *International Journal of Computer Science and Mobile Computing* 5:24–33.

Mirinejad, Hossein, and Tamer Inanc. 2015. "RBF Method for Optimal Control of Drug Administration in the Anemia of Hemodialysis Patients." *2015 41st Annual Northeast Biomedical Engineering Conference, NEBEC 2015*. doi: 10.1109/NEBEC.2015.7117180.

Panwong, Patcharaporn, and Natthakan Iam-On. 2016. "Predicting Transitional Interval of Kidney Disease Stages 3 to 5 Using Data Mining Method." *2016 2nd Asian Conference on Defence Technology, ACDT 2016* 145–50. doi: 10.1109/ACDT.2016.7437659.

Patil, Smitha. 2017. "A Literature Review on Prediction of Chronic Kidney Disease Using Data Mining Techniques." *International Journal of Science and Research* 6(3):2015–18.

Potharaju, Sai Prasad, and M. Sreedevi. 2016. "Ensembled Rule Based Classification Algorithms

for Predicting Imbalanced Kidney Disease Data." *Journal of Engineering Science and Technology Review* 9(5):201–7. doi: 10.25103/jestr.095.31.

S.Ramya, Dr. N. Radh. 2016. "Diagnosis of Chronic Kidney Disease Using Machine Learning Algorithms." *Technology* 4(2):2257–63.

Somji, Samina S., Pascal Ruggajo, and Sibtain Moledina. 2020. "Adequacy of Hemodialysis and Its Associated Factors among Patients Undergoing Chronic Hemodialysis in Dar Es Salaam , Tanzania." *Hindawi: International Journal of Nephrology* 2020.

Sung, Junne Ming, Chi Ting Su, Yu Tzu Chang, Yu Ru Su, Wei Chuan Tsai, Saprina P. H. Wang, Chun Shin Yang, Liang Miin Tsai, Jyh Hong Chen, and Yen Wen Liu. 2014. "Independent Value of Cardiac Troponin T and Left Ventricular Global Longitudinal Strain in Predicting All-Cause Mortality among Stable Hemodialysis Patients with Preserved Left Ventricular Ejection Fraction." *BioMed Research International* 2014. doi: 10.1155/2014/217290.

Yeh, Jinn Yi, Tai Hsi Wu, and Chuan Wei Tsao. 2011. "Using Data Mining Techniques to Predict Hospitalization of Hemodialysis Patients." *Decision Support Systems* 50(2):439–48. doi: 10.1016/j.dss.2010.11.001.