CHAPTER-43

PREDICTION OF CHRONIC KIDNEY DISEASE (CKD) USING A FUZZY INFERENCE SYSTEM

[1] HEMLATA SAHAY
[1] *Assistant Professor, Dept. of Mathematics, Dharam Samaj College, Aligarh*

[2] SHUBHNESH KUMAR GOYAL
[2] *Professor, Dept. of Mathematics, Dharam Samaj College, Aligarh*

Chapter ID: NSP/ICAAR-2023/A-43

ABSTRACT

Early diagnosis of CKD is vital in the fight to limit the number of people affected by the disease. As such, the goal of this research is to use the predictive model for CKD detection in MATLAB (fuzzy logic toolbox). The process of creating the model consists of five stages. At first, we established what tests – blood urea nitrogen, eGFR (estimated glomerular filtration rate), and serum creatinine – were used as input data. Second, min-max normalization processing was used to fuzzification of the inputs and outputs. After that, we built an inference engine and aggregated our rules. Finally, defuzzification of the predictive model outputs was used to examine each patient's CKD status. In conclusion, detecting CKD early is crucial for treating the condition effectively. A nephrologists or other specialist can then be consulted for further care before any major concerns arise.

Keywords: *fuzzy inference system, MATLAB, CKD, eGFR, blood urea nitrogen test*

INTRODUCTION

Large amounts of data with characteristics of variety, velocity, authenticity, and volume make healthcare big data analytics a necessity in the twenty-first century. Each second, this healthcare data is gathered from a wide range of global healthcare providers. All parties involved in healthcare, including providers, payers, patients, and management, stand to gain much from utilizing big data for analytical purposes. Chronic kidney disease is a global health issue that has to be addressed. There is a massive quantity of data being collected on CKD, and this may be used to inform effective decision making across the healthcare system to combat this disease.

Chronic kidney disease (CKD) is kidney damage caused by causes connected to modern living. Changes in health are also being noted now due to exposure to environmental changes, although many people are too preoccupied to notice them. Lack of water intake, smoking, poor food, lack of sleep, and other factors can all contribute to chronic kidney disease. Studies have also shown that diabetes is a major contributor to the development of renal failure. Unlike the vast majority of diseases, CKD is typically not

diagnosed until the very end, when kidney failure has already set in and treatment is extremely dangerous and costly.

This research tries to construct a method for estimating the danger of developing CKD, taking into account all of the symptoms and causes that contribute to the disease. Kidney disease phases will be characterised by their associated symptoms. A group of patients' medical histories can be categorised according to the various phases of renal disease. When used to patient classification, it facilitates the identification of CKD's most prominent features. When it comes to the most prominent characteristics, there are measures that may be taken to halt the development of CKD.

Feature selection is necessary for identifying the most important characteristics before any machine learning method can be applied. The most important characteristics are chosen with the help of a feature selection technique called random forest. The focus of this work is on the application of unsupervised learning strategies, such as neuro-fuzzy systems and clustering. These methods will classify a patient's renal condition into one of several possible categories. Loss of kidney function over time is the hallmark of chronic kidney disease (CKD). Kidneys are responsible for excreting excess fluid and waste products from the body through urination. As a result of incorporating fuzzy logic into the verification process, a condition can take on values other than true or false. It allows for more leeway in reasoning, which is especially helpful when dealing with probabilistic uncertainty.

In recent years, chronic kidney disease (CKD) has become a leading killer in the world worldwide. The prevalence of chronic kidney disease (CKD) is on the rise, affecting an estimated 800 million people worldwide [1] (World Health Organisation). The topic of how to detect CKD early enough to ease the strain on healthcare providers and institutions is of paramount importance. One viable option here is an intelligent health system designed to spot signs of CKD in its earliest stages. Increased efforts by numerous researchers towards better prevention and treatment of CKD are warranted in light of the disease's high prevalence and substantial negative impact. Early diagnosis and effective management are merely treatments to reduce mortality from CKD, which progresses slowly over time. Work centred on disease prediction is necessary to solve CKD issues. Early detection of CKD allows for more effective treatment of the disease.

Using ANN and SVM implementations in MATLAB, **Vijayarani and Dhayanand (2015)** classified CRD into four subtypes. The primary goal was to assess the relative accuracy and efficiency of different methods for predicting CKD. It was found that ANN networks were more accurate than SVM machines. **Chimwayi et al. (2017)** used hierarchical clustering to show that chronic renal disease and diabetes are strongly linked. In order to classify patients with chronic renal disease, **Oo Nway (2018)** employed a principal component analysis (PCA) based feature selection method and three rule based algorithms. In order to enhance our FIS with the ANFI system for the detection of chronic kidney disease using real-world data, **Zarandi and Abdolkarimzadeh (2018)** built a type-1 fuzzy inference system. **Adnan et al. (2019)** reviewed the components of CKD and used fuzzy logic to determine the relationships between the variables. Renal cancer diagnosis was the focus of **Singla et al. (2020)** proposed new multilayer fuzzy inference method. By applying the proposed expert system to modelling the progression of kidney cancer based on medical guidance, they were able to accurately offer an analysis of the results. To aid in the diagnosis of kidney cancer, **Nikita and Sadawarti (2020)** created a multi-layered fuzzy Sugeno model. The geographic distribution of chronic kidney disease was determined by Andoh et al. (2021) using the modification of diet for renal disease model adjusted for age, race, sex, and creatinine.

Using fuzzy and adaptive neural fuzzy inference systems, Murugesan et al. (2022) identified chronic kidney disease. Their primary goal is to improve the accuracy of medical diagnostics used to identify disease. Kumar et al. (2023) introduced a unique deep learning model for the recognition and prediction of renal illness by combining a fuzzy deep neural network.

EXPLANATION OF THE MEDICO-LEGAL TERMS EMPLOYED IN THE RESEARCH

Here, we define some of the medical jargon specific to chronic kidney disease.

1. **Blood urea nitrogen test-** Blood urea nitrogen (BUN) is a diagnostic test used to evaluate the concentration of urea nitrogen in the blood. As a byproduct of protein metabolism, the liver undergoes a process called the urea cycle, where urea is produced. Urea nitrogen concentrations in the blood of healthy adults should range from 2.1 to 7.1 mmol/L (6-20 mg/dL). Different reference ranges will be established by various laboratories since the assays employed can vary. In order to diagnose renal issues, this test is performed. This blood test is not as trusted as the creatinine test or the BUN/creatinine ratio test.
2. **Serum Creatinine Test-** Muscles convert creatine into creatinine as a waste product. Creatinine is a waste product that is normally carried by the blood to the kidneys and then excreted in the urine. A high blood level may be an indicator of impaired renal function. Creatinine levels typically range between 0.7 and 1.2 mg/dL in males and 0.5 and 1.0 mg/dL in females. Extremely high or low levels may point to renal disease or weakness, respectively. Creatinine levels may change due to medication use or pregnancy.
3. **EGFR (estimated glomerular filtration rate):** Kidney function can be measured via a test called the glomerular filtration rate (GFR). Glomeruli are microscopic filters in the kidneys that clean the blood of toxins and other unwanted substances. The filtration process's efficacy is measured by the GFR test. It is usual practise to measure eGFR before to initiating medical procedures or treatments that may have an impact on the kidneys in order to diagnose renal failure or other kidney issues.

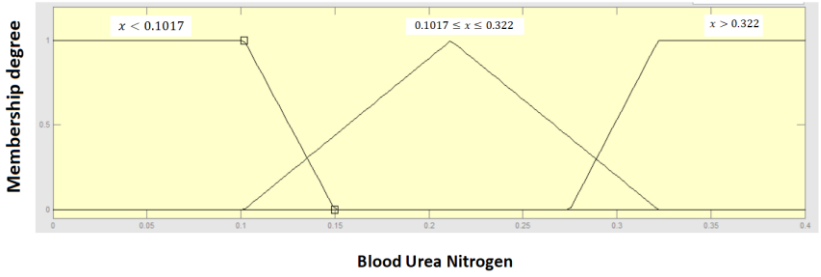
The glomerular filtration rate (GFR) is reported in volume/time units, such as milliliters/minute (mL/min)

$$GFR = \frac{Urine\ concentration \times Urine\ flow}{Plasma\ Concentration}$$

The kidneys are functioning at least 60% when the eGFR is over 60. In most cases, a higher value indicates healthier kidneys.

MEMBERSHIP PLOT FUNCTIONS OF INPUT AND OUTPUT VARIABLES

Figure 1: Membership plot function of blood urea nitrogen

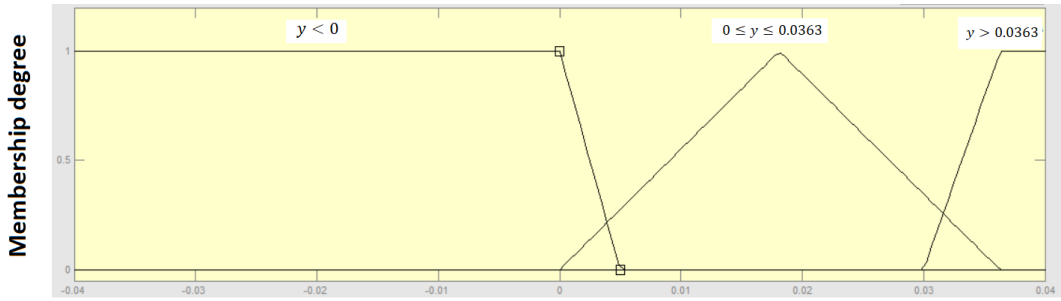


$$\mu_1(x) = \begin{Bmatrix} 1 & 0 \le x < 0.1017 \\ -20.7x + 3.10 & 0.1017 \le x \le 0.15 \end{Bmatrix}$$

$$\mu_2(x) = \begin{Bmatrix} 9.07x - 0.9232 & 0.1017 \le x < 0.21185 \\ -9.07x + 2.92 & 0.21185 \le x \le 0.322 \end{Bmatrix}$$

$$\mu_3(x) = \begin{Bmatrix} 9.07x - 0.9232 & 0.275 \le x \le 0.322 \\ -1 & x > 0.322 \end{Bmatrix}$$

Figure 2: Membership plot function of Serum creatinine test

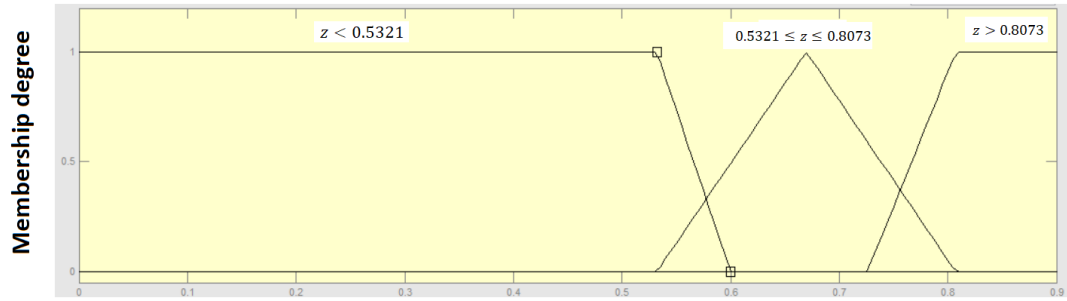


$$\mu_1(y) = \begin{Bmatrix} 1 & y < 0 \\ -200y + 1 & 0 \le y \le 0.005 \end{Bmatrix}$$

$$\mu_2(y) = \begin{Bmatrix} 55.9x & 0 \le y < 0.01815 \\ -55.09x + 2 & 0.01815 \le y \le 0.0363 \end{Bmatrix}$$

$$\mu_3(y) = \begin{Bmatrix} 158.73x - 4.76 & 0.03 \le y \le 0.0363 \\ 1 & y > 0.0363 \end{Bmatrix}$$

Figure 3: Membership plot function of eGFR

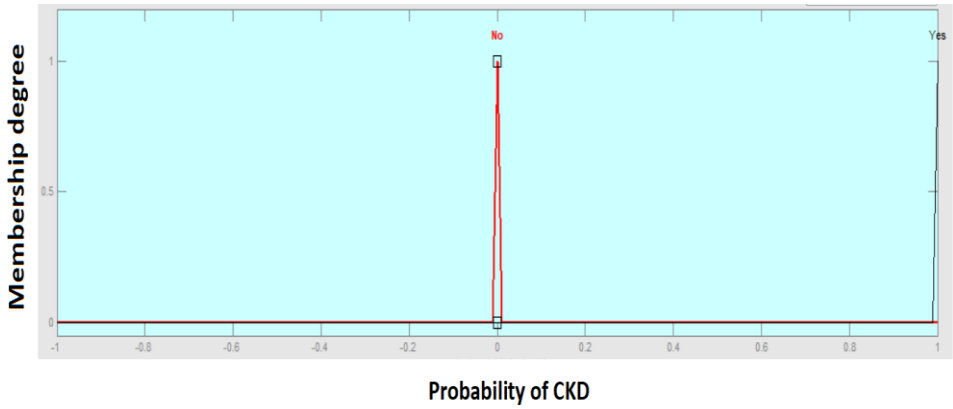


$$\mu_1(z) = \begin{Bmatrix} 1 & z < 0.5321 \\ -14.72z + 8.83 & 0.5321 \le z \le 0.6 \end{Bmatrix}$$

$$\mu_2(z) = \begin{Bmatrix} 7.26z - 3.86 & 0.5321 \le z < 0.6697 \\ -7.26z + 5.86 & 0.6697 \le z \le 0.8073 \end{Bmatrix}$$

$$\mu_3(z) = \begin{Bmatrix} 11.79z - 8.52 & 0.7225 \le z \le 0.8073 \\ 1 & z > 0.8073 \end{Bmatrix}$$

Figure 4: Membership plot function of Probability of CKD



$\mu_1(t) = \{No \quad t = 0\}$

$\mu_2(t) = \{Yes \quad t = 1\}$

4. Rule base: To establish the relationships between input and output, we have used the logical "AND" and "OR" operators in this section.

Figure 5: Rule base interface

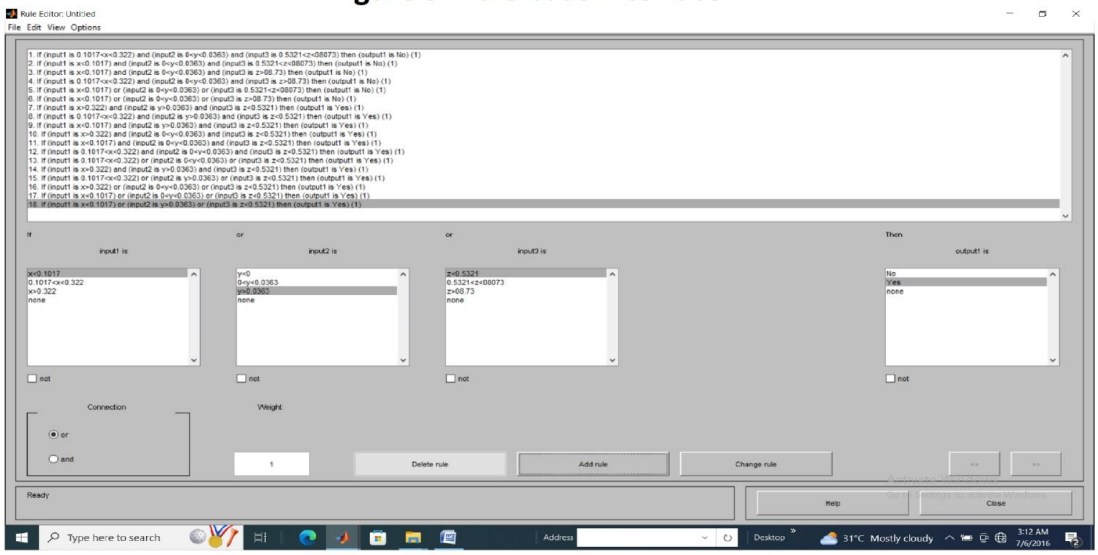


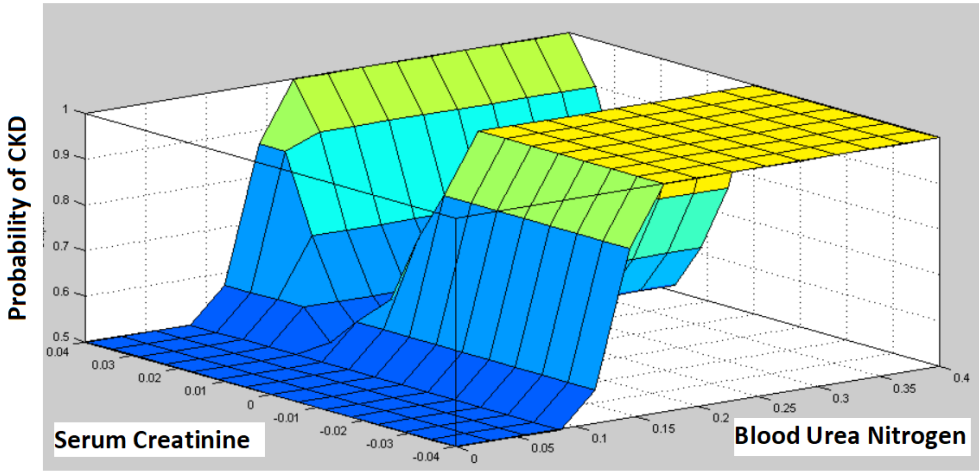
Table 1: Rule base of the FIS

Rule	Blood Urea Nitrogen	Operators	Serum Creatinine	Operators	eGFR	Probability of CKD
1.	$0.1017 \leq x \leq 0.3220$	AND	$0 \leq y \leq 0.0363$	AND	$0.5321 \leq z \leq 0.8073$	NO

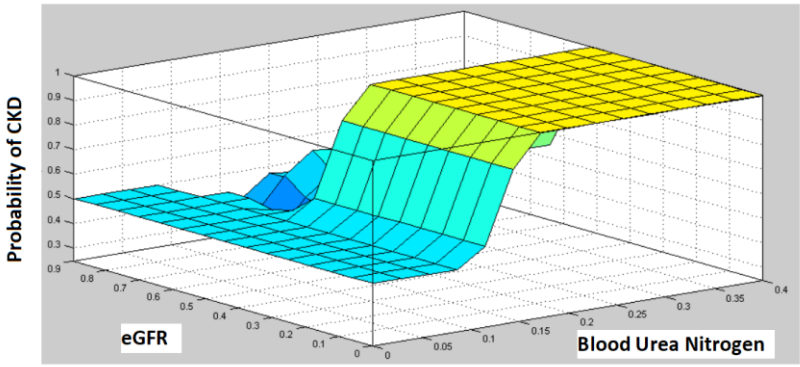
2.	$x < 0.1017$	AND	$0 \leq y \leq 0.0363$	AND	$0.5321 \leq z \leq 0.8073$	NO
3.	$x < 0.1017$	AND	$0 \leq y \leq 0.0363$	AND	$z \geq 0.8073$	NO
4.	$0.1017 \leq x \leq 0.3220$	AND	$0 \leq y \leq 0.0363$	AND	$z > 0.8073$	NO
5.	$x < 0.1017$	OR	$0 \leq y \leq 0.0363$	OR	$0.5321 \leq z \leq 0.8073$	NO
6.	$x < 0.1017$	OR	$0 \leq y \leq 0.0363$	OR	$z > 0.8073$	NO
7.	$x > 0.3220$	AND	$y > 0.0363$	AND	$z < 0.5321$	YES
8.	$0.1017 \leq x \leq 0.3220$	AND	$y > 0.0363$	AND	$z < 0.5321$	YES
9.	$x < 0.1017$	AND	$y > 0.0363$	AND	$z < 0.5321$	YES
10.	$x > 0.3220$	AND	$0 \leq y \leq 0.0363$	AND	$z < 0.5321$	YES
11.	$x < 0.1017$	AND	$0 \leq y \leq 0.0363$	AND	$z < 0.5321$	YES
12.	$0.1017 \leq x \leq 0.3220$	AND	$0 \leq y \leq 0.0363$	AND	$z < 0.5321$	YES
13.	$0.1017 \leq x \leq 0.3220$	OR	$0 \leq y \leq 0.0363$	OR	$z < 0.5321$	YES
14.	$x > 0.3220$	AND	$y > 0.0363$	AND	$z < 0.5321$	YES
15.	$0.1017 \leq x \leq 0.3220$	OR	$y > 0.0363$	OR	$z < 0.5321$	YES
16.	$x > 0.3220$	OR	$0 \leq y \leq 0.0363$	OR	$z < 0.5321$	YES
17.	$x < 0.1017$	OR	$0 \leq y \leq 0.0363$	OR	$z < 0.5321$	YES
18.	$x < 0.1017$	OR	$y > 0.0363$	OR	$z < 0.5321$	YES

RESULTS AND DISCUSSION

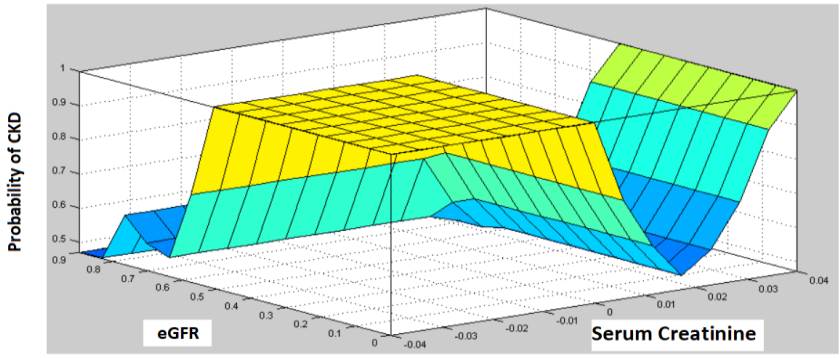
Graph 1: Surface plot of probability of CKD for blood urea nitrogen and serum creatinine

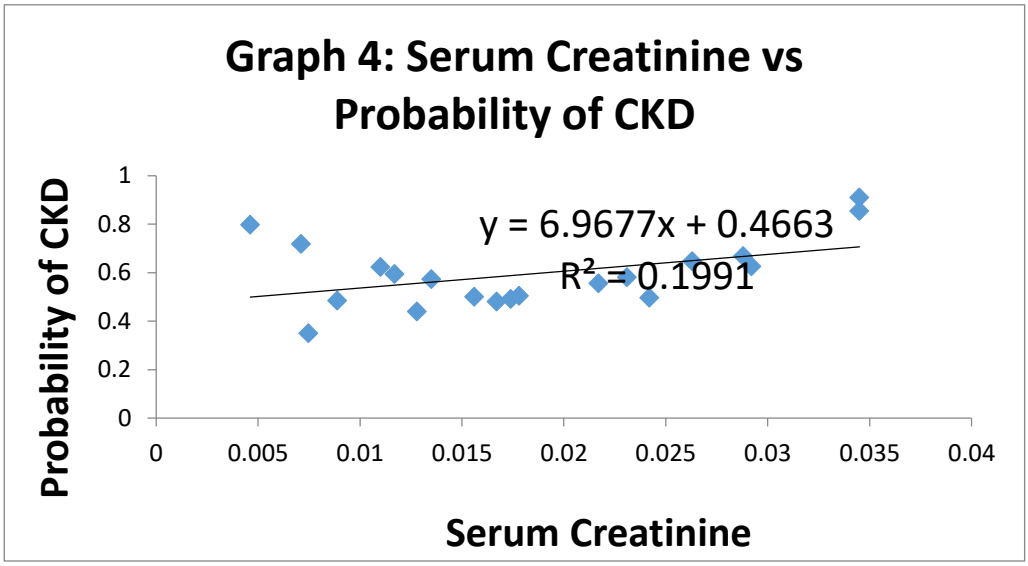


Graph 2: Surface plot of probability of CKD for blood urea nitrogen and eGFR



Graph 3: Surface plot of probability of CKD for serum createnine and eGFR





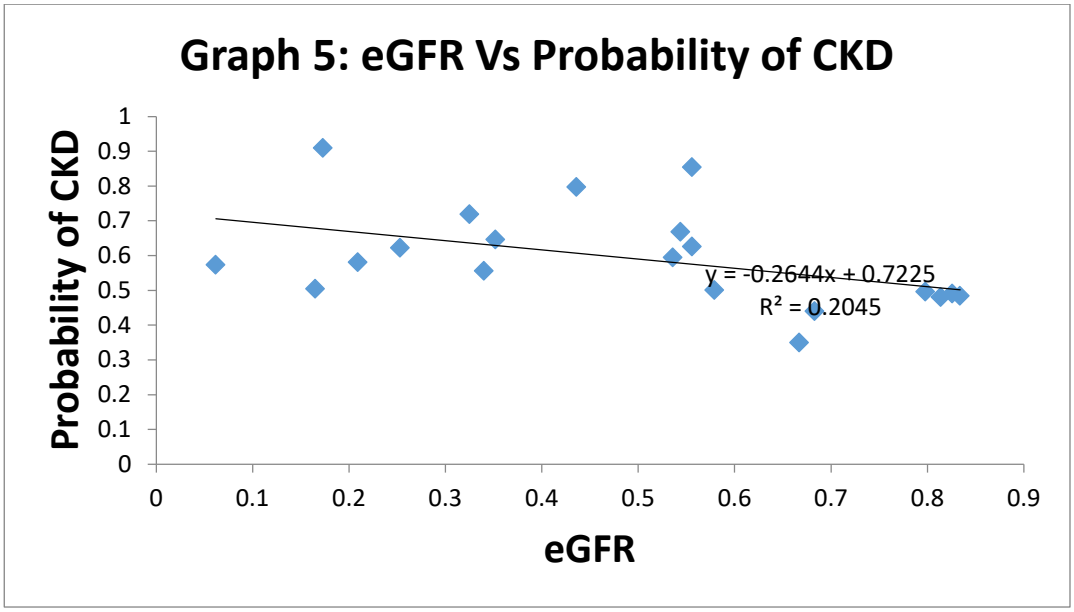


Table-2: Estimated values of probability of CKD for different input variables

S.No.	Blood Urea Nitrogen	Serum Creatinine	eGFR	Probability of CKD
1.	0.189	0.0178	0.165	0.504
2.	0.211	0.011	0.253	0.622
3.	0.259	0.0135	0.0617	0.573
4.	0.31	0.0231	0.209	0.58
5.	0.25	0.0263	0.352	0.646
6.	0.143	0.0117	0.536	0.594
7.	0.209	0.00462	0.436	0.797
8.	0.259	0.0128	0.683	0.439
9.	0.0684	0.0156	0.579	0.5
10.	0.198	0.0242	0.798	0.496
11.	0.0305	0.0345	0.173	0.909
12.	0.253	0.0288	0.544	0.668
13.	0.152	0.0167	0.814	0.48
14.	0.2	0.0217	0.34	0.555
15.	0.326	0.0345	0.556	0.854
16.	0.159	0.0292	0.556	0.625
17.	0.308	0.0174	0.826	0.49
18.	0.264	0.00747	0.667	0.349
19.	0.319	0.00889	0.834	0.484
20.	0.217	0.00711	0.325	0.718

The results of a MATLAB R2009 simulation are shown in the graphs (1, 2, and 3). Based on the chronic kidney disease data provided as input, these surfaces will display the associations and correlations between CKD function and features. It is currently challenging to accurately represent greater dimensions in output graphs using a fuzzy logic tool. Therefore, we could only compare two inputs to a single output, and that's why we did it. However, the system's functionality was not compromised, as the factors represented are all vital to the operation of the entire. The relationship between serum creatinine and blood urea nitrogen and the risk of developing chronic kidney disease is depicted in graph (1). Based on the results, it appears that both creatinine and blood urea, but especially blood nitrogen urea, have a role in determining the likelihood of CKD. The likelihood of identifying CKD is shown to be dependent on eGFR and blood urea nitrogen levels in graph (2). According to the results, a very high blood urea nitrogen level is more important than any other factor in predicting CKD. The relationships between eGFR and serum creatinine, which serve as predictors of CKD risk, are depicted in Graph (3). The results demonstrate that eGFR and serum creatinine can be used in conjunction with other risk factors like age and gender to predict the likelihood of developing CKD. The correlation between serum creatinine and eGFR, as well as a scatter diagram and trend line for the likelihood of CKD, are depicted in graphs (4) and (5), respectively. Serum creatinine is directly correlated to the risk of developing chronic kidney disease, as shown in figures (4). Graph (5) also shows that a higher eGFR is associated with a lower risk of developing CKD.

CONCLUDING REMARKS

Predicting chronic kidney failure and illness is a challenging problem, and numerous strategies have been investigated. Fuzzy logic was used in this investigation to aid in the detection of chronic renal failure and illness. Three tests, blood urea nitrogen, serum creatinine, and estimated glomerular filtration rate (eGFR), were used to demonstrate the proposed model for predicting CKD. Each patient's CKD status is now known thanks to the successful implementation of five sequential procedures. An inference system for the fuzzy inference engine was built using the significance of variables. The crisp value was normalised to a fuzzy value between 0 and 1 using the min-max method. After that, we fuzzed all the variable values and fed the resulting data into the inference engine. The inference engine's output of 13 rules was consolidated into a single set. Every patient's CKD status was revealed by the predictive model's results. In conclusion, it is hoped that nephrologists will be able to use this model to better identify patients with chronic kidney failure and related diseases. This technique also aids in cutting down on the high price tag of haemodialysis treatments and the number of avoidable deaths caused by a delay in diagnosis. In order to acquire a more representative sample size for assessment, future studies may take into account a larger number of patients. Furthermore, depending on the goals and constraints of the study, the number of indicators employed may be increased.

REFERENCES

1. *Adnan A.N.B.M., Zakaria N.I.B, Abdullah N.S.K.B, Acob S.Y. (2019): "Fuzzy logic: an application to detect chronic kidney disease and failure", Open International Journal of Informatics (OIJI), 7(2):118-126.*
2. *Andoh B., Adebayo A.F., Weyori B.A., Tannor E.K. (2021): "An adaptive neuro-fuzzy expert system for clinical staging of patients with chronic kidney disease", Research Square, 1-15.*
3. *Chimwayi K.B., Haris N., Caytiles R.D., Iyengar S.N. (2017): "Risk Level Prediction of Chronic Kidney Disease Using Neuro- Fuzzy and Hierarchical Clustering Algorithm (s)", International Journal of Multimedia and Ubiquitous Engineering, 12(8):23-36.*
4. *Fazel Zarandi M.H., Abdolkarimzadeh M. (2018): "Fuzzy rule based expert system to diagnose chronic kidney disease", Advances in Intelligent Systems and Computing, 648:323-328.*

5. *Kumar K., Pradeepa M., Mahdal M., Verma S., Raja Rao M.V.L.N., Ramesh J.V.N (2023): "A deep learning approach for kidney disease recognition and prediction through image processing", Applied Sciences, 13:1-14.*

6. *Murugesan G., Ahmed T.I., Bhola J., Shabaz M., Singla J., Rakhra M., More S., Samori I.A.(2022): "Fuzzy logic-based systems for the diagnosis of chronic kidney disease", BioMed Research International, Article ID 2653665, 15 pages.*

7. *Nikita, Sadawari H. (2020): "The Diagnosis of Renal Cancer using Sugeno Fuzzy Inference System", International Journal of Engineering Research & Technology (IJERT), 8(16): 160-165.*

8. *Oo Nway A. (2018): " Classification of chronic kidney disease (CKD) using rule based classifier and PCA", International Journal of Management, Technology And Engineering, 8(7): 728-732.*

9. *Singla N., Sadawarti H., Singla J., Kaur B. (2020): "Development of multilayer fuzzy inference system for diagnosis of renal cancer", Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 39(1):885-898.*

10. *Vijayarani S., Dhayanand S. (2015): "Kidney disease prediction using SVM and ANN algorithms", International Journal of Computing and Business Research (IJCBR), 6(2):2229-6166.*