International Journal of Innovative Mathematics, Statistics & Energy Policies
12(4):40-49, Oct-Dec., 2024

 www.seahipublications.org **ISSN: 2467-852X**

DOI:10.5281/zenodo.14016912

Predicting Prevalence Of Chronic Kidney Disease Using Possibilistic Fuzzy Linear Regression

Danjuma Abdulazeez, Kazeem E. Lasisi, Bello A. Rasheed, Musa Usman Bawa, Abdulmalik Mohammed & Kamalu Ibrahim Balansana.

Department of Statistics, Federal Polytechnic Damaturu, Yobe State, Nigeria
Department of Statistics, Abubakar Tafawa Balewa University, Bauchi State, Nigeria
Department of Mathematical Science, Gombe State University, Gombe, Nigeria
Department of Statistics, Abubakar Tafawa Balewa University, Bauchi, Nigeria
Department of Computer/Mathematics, C.O.E Dutsen Tanshi Bauchi, Bauchi State, Nigeria
Department of Statistics, Modibbo Adamawa University, Yola, Adamawa, Nigeria

Corresponding Author: Danjuma Abdulazeez, Education
Email: abdulazeezdanjuma700@gmail.com

ABSTRACT
Chronic kidney disease (CKD) is a prevalent health issue worldwide, with significant implications for both individuals and healthcare systems. Early prediction of CKD can lead to a better management. In this paper, we propose an approach using Possibilistic Fuzzy Regression (PFLR) to predict the prevalence of CKD. By applying Possibilistic Fuzzy regression and Logistic regression model. A data on Chronic Kidney Disease and it risk factors were collected from January, 2023 to December, 2023, and the considered variables for prediction of CKD include Age, Blood pressure, Diabetes Mellitus, Hypertension, Anemia and Edema. The result of the study showed that both models under consideration can be used to fit and predict prevalence of CKD data. But the Logistic regression model cannot handle the Crisp and Vagueness nature of the CKD data and also cannot cope with presence of outliers which may affect the regression coefficients. The study conclude that Possibilistic fuzzy regression handles the effect of outliers more efficiently. Therefore, the best model for predicting prevalence of chronic Kidney diseases as an alternative to logistic regression is Possibilistic Fuzzy regression model, which provides valuable insights for healthcare providers and policymakers in designing targeted interventions for CKD prevention and management.
Keywords: Chronic Kidney Disease, Possibilistic Fuzzy Regression Model, Outliers, Crisp data

INTRODUCTION
Chronic kidney disease (CKD) is defined as a reduction in glomerular filtration rate (GFR) below 60 ml/min per 1.73 m2 or an indication of abnormalities in the kidney structure of those who had known CKD (SF Gilbert., et al 2022). In patients with CKD, excess fluid and waste materials from the blood remain in the body and are not excreted from the body, leading to severe health problems (Stats F. 2017). Chronic kidney disease is a silent killer. A person can lose up to 90% of kidney function before

experiencing any signs or symptoms. Most people have no symptoms until CKD is advanced. Untreated kidney failure is life-threatening, so the early detection of falling kidney function is crucial because it allows for suitable treatment before the kidney is damaged (Kaze et al 2018). Chronic kidney disease is an important and common public health problem in NCDs. It affects as much as 10%–15% of the world's population (Shiferaw F., et al 2018). Globally, CKD represents a major public health issue that can consume substantial financial and social resources (Lancet. 2017). Currently, more than 2 million people globally receive treatment with dialysis or kidney transplants to stay alive; of them, 20% are treated in 100 low-income countries, which make up half of the world's population (Bomback AS 2011). In 2017, the global burden of disease (GBD) reported that CKD ranks the 16th leading cause of premature death and is estimated to be the fifth leading cause of premature death by 2040. The pooled prevalence of CKD in Africa and sub-Saharan Africa was 10% and 14%, respectively (Levey AS, 2009). Kidney disease is associated with a great economic burden. High-income countries typically spend more than 2%–3% of their annual healthcare budget on the treatment of end-stage renal disease (ESRD). Next to cardiovascular complications, CKD is a serious public health issue among people with type 2 diabetes (T2DM). Patients with type 2 diabetes are more likely to require recurrent hospital stays and are at greater risk of increased mortality (Ademuyiwa A., et al 2021). CKD is a well-recognized and serious complication of diabetes, and diabetes is one of the most common causes of CKD, with up to 44% of patients with CKD affected by diabetes (Li S, et al. 2011). The prevalence of CKD among patients with hypertension (HTN) was 21.1% in Tigray (Bahrey D., et al 2019), 46.9% in Gahanna (Osafo C., et al 2011), and 17.6% in Northwest Ethiopia referral hospitals (Hunegnaw A,et al 2021). Geographically, the prevalence of CKD linked to human immunodeficiency virus (HIV) varies, ranging between 2% and 38% due to genetic variation, the start of anti-retroviral therapy (ART), and the various definitions of CKD in each region (Ekrikpo UE., et al 2018). The global prevalence of chronic kidney disease among persons living with HIV (PLHIV) is 6.4%. This prevalence varies by continent: 7.9% in Africa, 7.1% in North America, 5.7% in Asia, and 3.7% in Europe (Ekrikpo UE., et al 2018). CKD if detected at its early stage, avoiding all the unpredictable, because its symptoms develop slowly, while some people have no symptoms at all. Whereas the symptoms are not specific to the disease. Hence fuzzy linear regression can be helpful in such cases for predicting those patient with CKD or not. Glomerular Filtration Rate (GFR) is a test used to check how well the kidneys are working. Specifically, it estimates how much blood passes through the glomeruli each minute. Glomeruli are the tiny filters in the kidneys that filter waste from the blood (Kore C., et al 2018). It can be calculated from the results of your blood creatinine, age, race, gender, and other factors. The earlier the disease is detected the better one learn how to take control and make any change to certain lifestyle habit in order to possibly slow the progression of your condition.

Logistic regression analysis is one of the famous non-linear methods used to model the binary response variable based on ordinary explanatory variables. This method is particularly appropriate for models involving disease states (diseased/healthy), patient survival (alive/dead) and decision making (yes/no). So, it is widely used in studies in the health sciences (Bagley, 2001). But on the other hand, the logistic regression cannot cope with vague data because it does not follow random error patterns, likewise it does not accommodate low sample size. In fact, non-precise or vague (fuzzy) observations, which occur frequently in practice, may cause all other difficulties in the logistic model, that may lead to inefficient result. Ali *et al* (2023), determined the Prevalence of chronic kidney disease and associated factors among patients with underlying chronic disease at Dessie Referral Hospital, East Amhara Region, Ethiopia using a Binary logistic regression model. So, looking at the nature of CKD data which is vague(fuzzy), the binary logistic regression would be inefficient. Therefore, this study intends to improve on their work by using a possibilistic fuzzy linear regression which will be able to describe the relation between crisp inputs-fuzzy binary output observations through the possibilistic fuzzy linear regression model. In fact, non-precise or vague observations, which occur frequently in practice, may cause all other difficulties. Consider clinical researches as an example; for some diseases, there are no biological examinations and the disease is diagnosed by some defined and wholly accepted criteria. Regression analysis is a powerful and comprehensive methodology to model the relationships between a response variable, called the dependent variable, and one or more explanatory variables called independent variables. By the classical statistical technique, the observations, either the response variable or the explanatory variables, are required to follow certain probability distributions (Kao and Chyu, 2002). In addition, deviations between the observed and the estimated values are assumed to be due to random errors. But, by the fuzzy observations, the distribution assumptions do not hold. Also, in this case, one cannot assign all the uncertainty of model to the randomness aspect of variables. In the analysis of fuzzy data, the study of the relationship between a set of (crisp or fuzzy) independent variables and one or more dependent fuzzy variables is an important issue. Such a relationship can be studied, for instance, through fuzzy linear regression techniques (Urso and Gastaldi, 2002). Fuzzy regression analysis is an extension of (or an alternative for) the classical regression analysis in which some elements of the model are represented by fuzzy numbers. The uncertainty in this type of regression model becomes fuzziness, not randomness. This aspect of uncertainty is called ‘‘possibility’’ (Zimmerman, 1991). The article by Tanaka et al. (Tanaka *et al*, 1982) is probably the first research on fuzzy linear regression. Fuzzy regression methods have been successfully applied to various problems such as forecasting (Tseng and Tseng, 2002) and engineering. There are two general categories of fuzzy regression analysis (Dom *et al*,2007); the first is a possibilistic regression analysis which is based on possibility concepts. Possibilistic regression analysis uses a fuzzy linear system as a regression model whereby the total vagueness of the estimated values for the dependent variables is minimized. It was first proposed by Tanaka et al. The second category of fuzzy regression analysis adopts the Fuzzy Least Squares Method (FLSM) for minimizing errors between the given outputs and the estimated outputs. The research adopted a Secondary data of (200) patients was collected from Yobe State University Teaching Hospital Damaturu Nigeria, covering January, 2023 to December, 2023. The considered variables for the prediction of prevalence of CKD include Age (AGE), Blood pressure (BP), Diabetes Mellitus (DM), Hypertension (HTN), Anemia (A) and Edema (E) respectively.

MATERIAL AND METHOD

The Fuzzy Linear Regression Model

In traditional statistical inference, as we know, the regression models are used frequently in the researches of the relations among several variables in a system. Observing some of the variables, we can make estimates and predictions for the others. If a system under consideration is not governed by random variables and/or crisp observation but is governed by possibility variables and/or imprecise observation, it is more natural to seek a fuzzy regression analysis for such a system.

Fuzzy linear regression is a fuzzy type of classical regression analysis in which some elements of the model are represented by fuzzy numbers. Fuzzy regression analysis was first proposed by Tanaka et al. (1980, 1982) where a fuzzy linear system is used as a regression model, (see also Tanaka, 1987; Tanaka et al.,1987). They consider a regression model in which the relation of the variables are subject to fuzziness, i.e., the model with crisp input and fuzzy parameters. Their approaches are developed in several ways, see for example, Peters (1994), Luczynski and Matloka (1995), Tanaka et al. (1995), Tanaka and Lee (1999), and Yen et al. (1999). The functional relationship between the response and explanatory variables as reported in Tanaka et al. (1982) is presented as follows:

$$\tilde{Y} = \tilde{A}_0 + \widetilde{A_1}x_1 + \widetilde{A_2}x_2 + \cdots \ldots + \widetilde{A_p}x_p \qquad (1)$$

$$OR$$

$$\hat{Y} = \tilde{A}_0 + \widetilde{A_1}\text{Age} + \widetilde{A_2}\text{BP} + \widetilde{A_3}\text{HTN} + \widetilde{A_4}\text{DM} + \widetilde{A_5}\text{EDEMA} + \widetilde{A_6}\text{ANAEMIA}.$$

In matrix form;

$$\tilde{Y} = \tilde{A}X \qquad (2)$$

Where: $\tilde{Y}$ is the fuzzy output, X $=(x_{0i}, x_{1i,} x_{2i,} \ \ldots \ldots x_{pi})^T$,

p-dimensional crisp input vector, $\tilde{A} = (\tilde{A}_0, \widetilde{A_1}, \widetilde{A_2}, \ldots \ldots \widetilde{A_p})^T$, fuzzy vector of coefficients presented in the form of a symmetric triangular fuzzy number denoted by $\tilde{A}_j = [c_j, w_j]$ respectively, $c_j\ and\ w_j$ are its centre and width, while x_{0i} = 1.

The Possibilistic method for fuzzy data

Tanaka modified his first model in 1987 and 1989(H. Tanaka, 1987 and Tanaka at al., 1989). The total "fuzziness" of the parameters (sum of) was minimized in the first model. On the contrary, the second model try to minimize the total fuzziness of the model. That model is called Possibilistic Fuzzy Linear Regression (PFLR).

$$z = \sum_{i=1}^{N} \left(c_0|x_0| + c_1|x_{1i}| + c_2|x_{2i}| + \cdots + c_n|x_{in}|\right) \quad (4)$$

Subject to

$$\alpha^t x_1 + (1-H)c^t|x_i| \geq y_i + (1-H)e_i$$
$$\alpha^t x_1 + (1-H)c^t|x_i| \leq y_i + (1-H)e_i$$

$$c_i \geq 0 \text{ for } i = 1,2,\dots,N$$

$$n\ is\ the\ number\ of\ individual\ variables\ and\ N\ is\ number\ of\ observations$$

Logistic regression model

The multiple binary logistics regression model is giving as follows:

$$\pi(X) = \frac{\exp(\beta 0 + \beta 1X1 + \cdots + \beta kXk)}{1 + \exp(\beta 0 + \beta 1X1 + \cdots + \beta kXk)}$$

$$= \frac{\exp(X\beta)}{1 + \exp(X\beta)}$$

$$= \frac{1}{1 + \exp(-X\beta)}$$

Where here π denotes a probability and not the irrational number 3.14…

Π is the probability that an observation is in a specified category of the binary Y variable, generally called the "success probability". We notice that the model describes the probability of an event happening as a function of *X* variables. For instance, it may provide estimates of the probability that an older person has heart disease. With the logistic model, estimate of π from equations like the one above will always be between 0 and1 the reasons are: The numerator $exp^{(\beta_0+\beta_1+\beta_{1x1}+\cdots+\beta_{kxk})}$ must be positive, because it is power of a positive value (e). The denominator of the model is (1+numerator), so the answer will always be less than 1. With one X variable, the theoretical model for π has an elongated *"S"* shape (or sigmoidal shape) with asymptotes at *0 and 1*, although in sample estimate we may not see this *"S"* shape if the range of *X* variable is limited.

For a sample of size n, the likelihood for a binary logistic regression is given by:

$$L(\beta; y, X) = \prod_{i=1}^{n} \pi i^{yi} (1 - \pi i)^{1-yi} = \prod_{i=1}^{n} \left(\frac{\exp(X\beta)}{1+\exp(X\beta)}\right)^{yi} \left(\frac{1}{1+\exp(X\beta)}\right)^{1-yi} \quad (5)$$

Model Specification

Taking chronic Kidney Disease as the dependent variable and Age (AGE), Blood pressure (BP), Diabetes Mellitus (DM), Hypertension (HTN), Anemia (A) and Edema (E) respectively as independent variables, the expected model will assume the form: Where $\beta_i ' s$ are estimated coefficients of the regression model such that, , *is the baseline constant for the regression model* β=relative risks, X_1 =Age, X_2 = Blood Pressure, X_3 = Diabetes Mellitus, X_4= Hypertension, X_5 = Anemia, X_6 = Edema.

RESULT AND DISCUSSION

Analysis

Data were analyzed using the R package version 4.4.0. The comparison among the models was done using cross validation statistical technique also known as MSE and RMSE. The MSE and RMSE of an estimator is one of many ways to quantify the difference between values implied by an estimator and the true values of the estimated quantity. This comparison was conducted to find the best and effective model for predicting the high risk factors for causing chronic Kidney disease in Nigeria, that would help the Medical researchers in getting accurate estimates of factors causing chronic Kidney disease.

Box plot of CKD prevalence determinant factors

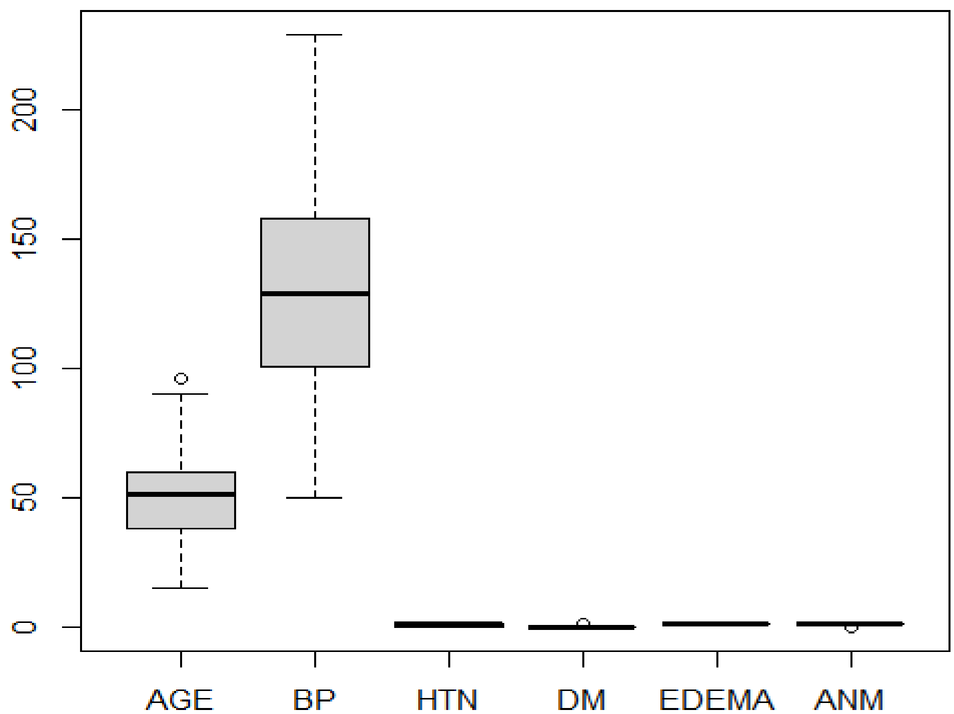

Figure 1: Boxplot of CKD data

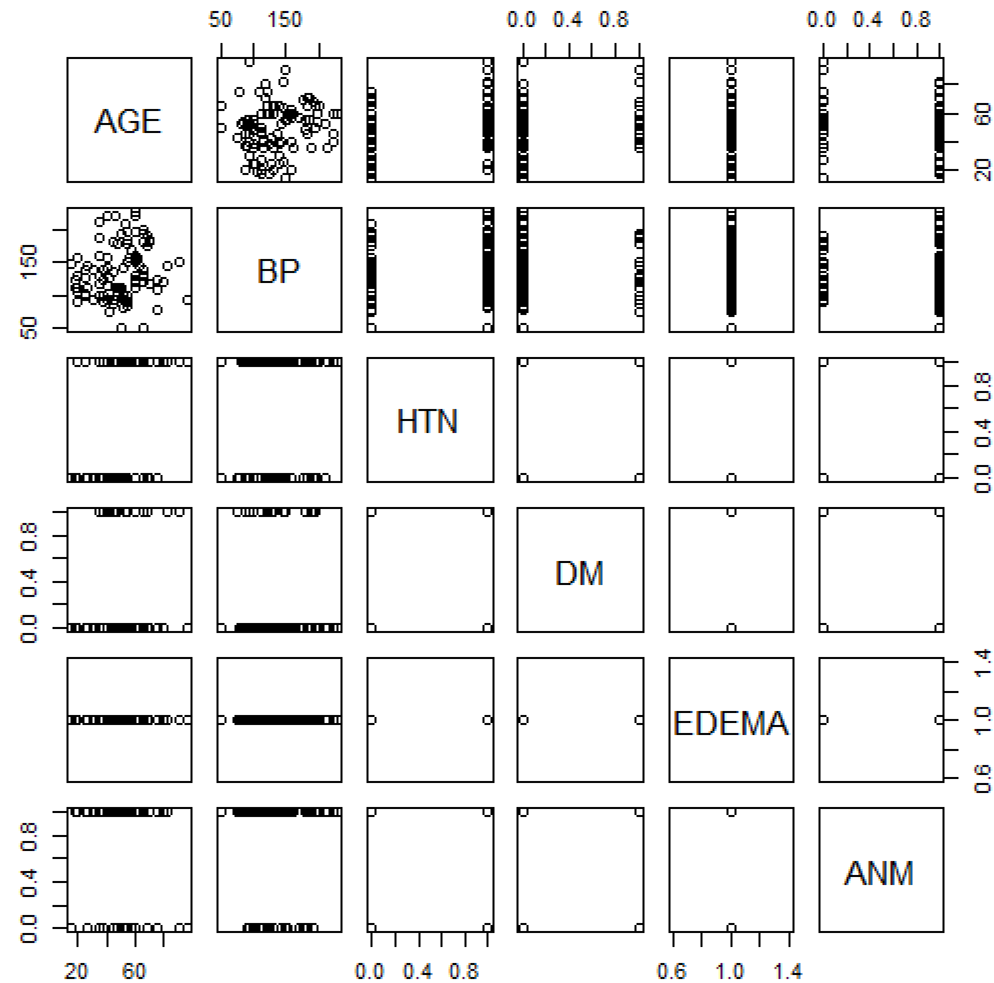

Figure 2: Boxplot of CKD data

Figure 1 and 2 above, displayed the boxplots of the CKD Prevalence determinants factors accordingly. Looking at figure 1, we can observe that only the variable BP found to be approximately symmetric. Age is skewed to left and also contain an outlier. Likewise, the variables DM and ANM also contain outliers as we can see points outside the whiskers.

Table 1. Possibilistic Fuzzy regression results of CKD data.

Fuzzy Parameter	Center α_i	Width c_i
Constant	0.5245	0
AGE	0.6521	0.9213
BP	1.7921	0
HTN	1.0171	0
DM	5.8210	0
EDEMA	-0.8112	0
ANAEMIA	7.1240	0
MSE	**0.2137**	
RMSE	**0.7123**	

Source: Authors' computation aided by R package v 4.4.0.

Table 1 displays the Possibilistic Fuzzy regression result with Center α_i and Width c_i values for the fuzzy parameters with Chronic Kidney Disease (CKD) as dependent variables against six independent variables. Looking at the table above, we can see that Anemia (A) found to have the highest fuzzy mean value (7.1240) followed by Diabetes Mellitus (DM) with fuzzy mean value of (5.8210). Out of the Six variables its only Age Width found to be non-zero (0.9213) which represents the level of the Parameters fuzziness. Likewise, the variable EDEMA found to be negative, meaning that the Chronic Kidney disease (CKD) is inversely proportional to EDEMA disease.
The following is the estimated fuzzy linear regression model for predicting Chronic Kidney disease
$\widehat{CKD}$= 0.5245 + 0.6521Age + 1.7921BP + 1.0171HTN + 5.8210 DM -0.8112EDEMA + 7.1240ANAEMIA.

Table 2: Ordinary logistic regression results of CKD data.

Variables	Coefficient	S.E	t-value	p-value
(Intercept)	-2.2033	0.2449	-8.996	2e-16 ***
AGE	0.4897	0.2203	2.676	0.4744
BP	1.6041	0.2432	7.006	2.46e-12 ***
HTN	0.6669	0.2228	4.340	1.43e-05 ***
DM	- 0.1058	0.2926	0.362	0.17825
EDEMA	0.5024	0.2555	3.532	0.20441
ANAEMIA	0.4417	0.5124	0.667	0.71771
MSE	**0.3066**			
RMS	**0.5537**			

Source: Authors' computation aided by R package v4.4.0. Note: (***) denote significance at 1%.

Table: 2. Presents the results of a logistic regression model predicting a binary outcome (CKD) based on several predictor variables. The table shows the estimated coefficients, standard errors, t-values, and p-values for each predictor, Age (AGE), Blood pressure (BP), Diabetes Mellitus (DM), Hypertension (HTN), Anemia (A) and Edema (E) are the categorical variables. The model's performance indicated by MSE of value (**0.3066**) and RMSE value of (**0.5537**) accordingly. The coefficients of the variables found all to be positive with the exception of Diabetes Mellitus (DM) which found to be negative (- 0.1058). Blood pressure (BP) and Hypertension (HTN) variables are the only variables found to be significant from the logistic result above.

Table 3: Comparing the performances of Possibilistic fuzzy regression and ordinary logistic regression models

Methods	MSE	RMS	
Logistic	0.3066	0.5537	
PFLR	0.2137	0.4623	

Source: Authors' computation aided by R package v 4.4.0

DISCUSSION OF RESULTS

Table 1 and 2 showed the results of Possibilistic Fuzzy regression and logistic model in predicting prevalence of chronic Kidney disease (CKD) in Nigeria respectively. While Table 3 displayed a comparison between the two models for predicting prevalence of chronic Kidney disease (CKD) in Nigeria using MSE and RMSE as performance measurement criteria. Looking at the table 3 result, we can see that the Possibilistic Fuzzy regression model (PFLR) found to have the lowest Mean Square error (MSE) and Root mean square error values of (0.2137) and (0.4623). Hence, we can notice from the result above that the Possibilistic Fuzzy regression models perform better than the ordinary logistic regression models. This is because of the Fuzzy nature of the CKD data that involves Crips and vague which cannot be captured well by the logistic regression model.

CONCLUSIONS

The result of the study showed that, both the two models under consideration, that its Logistic regression model and Possibilistic Fuzzy regression (PFLR) approaches can be used to fit and predicted prevalence of Chronic Kidney Disease (CKD) data. But the Logistic regression model cannot handle the Crips and Vagueness nature of the CKD data at the same time it cannot cope with presence of outliers which may affect the regression coefficients. The study also conclude that the Possibilistic fuzzy regression handles the effect of outliers more efficiently. Therefore, the study recommended that the best model to be used for predicting prevalence of chronic Kidney disease as an alternative to logistic regression is Possibilistic Fuzzy regression model (PFLR). The high-risk factors of CKD can be detected with greater precision by using the fuzzy linear regression method, which provides a better prediction of imprecise data than Logistic regression model, which provides valuable insights for healthcare providers and policymakers in designing targeted interventions for CKD prevention and management. The study found that ANAEMIA was the most impactful factor in causing chronic Kidney disease, as it has the highest fuzzy mean parameter in the model, followed by Diabetes Mellitus.

REFERENCES

Abu Syed *et al*, (2020). Identification of High Leverage Points in Linear Functional Relationship Model. *A Pakistan journal of Statistics and Operation Research*.Vol.16 No.3, pp 491-500.

A. Celmins, (1987) least squares model fitting to fuzzy vector data, Fuzzy Sets and Systems 22 (260–269).

Ali A, Mulatu K, Feleke S.F and Wassie G.T (2023). Prevalence of chronic kidney disease andassociated factors among patients with underlying chronic disease at Dessie Referral Hospital, East Amhara Region, Ethiopia. Front. Epidemiol. 3:1154522. doi: 10.3389/fepid.2023.1154522.

Bahrey, D., Gabremedhn, G., Mariye, T., Girmay, A., Aberhe, W. Hika, A. et al. (2019) Prevalence and associated factors of chronic kidney disease among adult hypertensive patients in Tigray teaching hospitals. *BMC Res Notes.*12, 562.

Begashaw, A.B and Yohannes, Y.B (2020). Review of outlier detection and identifying usin robust regression. *International Journal of Systems Science and applied mathematics.* 5(1): 4- 11.

Biccard, B. M., Gopalan P. D., Miller M, Michell, W. L., Thomson, D., Ademuyiwa, A. et al. (2021) Patient care and clinical outcomes for patients with COVID-19 infection admitted to African high-care or intensive care units (ACCCOS). *Lancet* 397(10288), 1885–1894.

C. Kao, C.L. Chyu, A fuzzy linear regression model with better explanatory power, Fuzzy Sets and Systems 126 (2002) 401–409.

Chen L.H. and Hsueh C.C. (2009) Fuzzy Regression Models Using the Least-Squares Method Basedon the Concept of Distance, *IEEE Transactions On Fuzzy Systems*, *17*, 6.

Chen Y.S. (2001). Outliers detection and confidence interval modification in fuzzy regression, *Fuzzy Sets and Systems*, *119,* 259-272.

Coppi R., D'Urso P. (2006) Giordani P., Santoro A., Least squares estimation of a linear regression model with LR fuzzy response, *Computational Statistics and Data Analysis*, *51* ,267 – 286.

Draper, N. R. & Smith, H. (1998). **Applied Regression Analysis**, Third Edition, Wiley Interscience Publication, United States, 1998.

D'Urso P., Gastaldi T. (2000). A least-squares approach to fuzzy linear regression analysis. *Computational Statistics and Data Analysis, 34,* 427-440.

Dwyer, J. P., Parving, H. H., Hunsicker, L. G., Ravid, M., Remuzzi, G., Lewis, J. B. et al. (2012) Renal dysfunction in the presence of normoalbuminuria in type 2 diabetes, Cardiorenal Med. 2(1), 1–10.

Ekrikpo, U. E., Kengne, A. P., Bello, A. K., Effa, E. E., Noubiap, J. J., Salako, B. L., et al. (2018) Chronic kidney disease in the global adult HIV-infected population. Fuzzy Logic Toolbox™, MATLAB, MathWorks.

F.M. Tseng, G.H. Tzeng, A fuzzy seasonal ARIMA model for forecasting, Fuzzy Sets and Systems 126 (2002) 367–376.

Gilbert, S. F., Weiner, D. E. (2022) National kidney foundation primer on kidney diseases, E-Book. Elsevier.

Göőz, M. Chronic kidney disease. (2012) BoD–Books on Demand.

H. Tanaka, S. Uejima, K. Asai, Linear regression analysis with fuzzy model, IEEE Transactions on Systems, Man and Cybernetics 12 (1982) 903–907.

Hunegnaw, A., Mekonnen, H. S., Techane, M. A., Agegnehu, C. D. (2020) Prevalence and associated factors of chronic kidney disease among adult hypertensive patients at Northwest Amhara Referral Hospitals, Northwest Ethiopia,

H.J. Zimmerman, Fuzzy Set Theory and its Applications, Kluwer Academic, Boston, 1991.

Hung W.L., Yang M.S. (2006) An omission approach for detecting outliers in fuzzy regression models, *Fuzzy Sets and Systems, 157* (23), 3109 – 3122.

IHME. The Lancet. Global burden of disease study. (2017). Kao C., Chyu C.L. (2002) A fuzzy linear regression model with better explanatory power, *Fuzzy Sets and Systems*, *126* 401 – 409.

Kaze, A. D., Ilori, T., Jaar, B, G., Echouffo. (2018) Burden of chronic kidney disease on the African continent a systematic review and meta-analysis. BMC Nephrol. 19(1), 1–11.

Lee E.S., Chang P.T. (1994) Fuzzy linear regression analysis with spread unconstrained in sign, *Comp. Math. Appl.*, *28*(4), 61-70.

Levey, A. S., Stevens, L. A., Schmid, C. H., Zhang, Y., Castro, Feldman, H. I. et al. (2009) A new equation to estimate glomerular filtration rate. Ann Intern Med. 150 (9), 604–612.

Levin, A., Tonelli, M., Bonventre, J., Coresh, J., Donner, J, A., Fogo, A., B. et al. (2017) Global kidney health and beyond, a roadmap for closing gaps in care, research and policy. Lancet 390(10105),1888–1917.

Lu J., Wang R. (2009) An enhanced fuzzy linear regression model with more flexible spreads, *Fuzzy Sets and Systems, 160,* 2505 – 2523. 43

McClave J.T. & Sincich T. (2009). *Statistics* (11th edition). Pearson Education Inc.,

McFarlane, S., I, McCullough, P., A, Sowers, J., R, Soe, K, Chen, S., C, Li, S. et al. (2011) Comparison of the CKD Epidemiology Collaboration (CKD-EPI) and Modification of Diet in Renal Disease (MDRD) study equations, prevalence of and risk factors for diabetes mellitus in CKD in the Kidney Early Evaluation Program (KEEP). Am J Kidney Dis. 57(3), S24–S31.

Myung I.J. (2003). Tutorial on maximum likelihood estimation, *Journal of Mathematical Psychology, 47*, 90–103

Nasrabadi M.M., Nasrabadi E. (2004). A mathematical-programming approach to fuzzy linear regression analysis, *Applied Mathematics and Computation, 155,* (3) 873 – 881.

Nasrabadi M.M., Nasrabadi E., Nasrabady A.R. (2005). Fuzzy linear regression analysis: a multi-objective programming approach, *Applied Mathematics and Computation, 163*, 245 – 251.

National Kidney Foundation. (2015) World kidney day. Chronic kidney disease. (2015). Available at:h ttp://www.worldkidneyday.org/faqs/chronic-kidney-disease (2015) 382:9888.

Neter J., & Kunter M.H., & Nachtsheim C. J., & Wasserman W.(1996) *Applied linear regression models*, (6th edition), Irwin.

Osafo, C., Mate-Kole, M., Affram, K., Adu, D. (2011) Prevalence of chronic kidney disease in hypertensive patients in Ghana. Ren Fail 33(4), 388–392.

Özelkan E.C., Duckstein L. (2000). Multi-objective f uzzy regression: a general framework, *Computers & Operations Research, 27*, 635 – 652.

Peters G. (1994). Fuzzy linear regression with fuzzy intervals, *Fuzzy Sets and Systems, 63*, 45-55.

P. D'Urso, T. Gastaldi, A least-squares approach to fuzzy linear regression analysis, Computational Statistics & Data Analysis 34 (2000) 427–440.

P. Diamond, least squares fitting of several fuzzy variables, in: Proc. of the Second IFSA Congress, Tokyo, 1987, pp. 20–25.

R.M. Dom, R. Zain, S. Abdul Kareem, B. Abidin, An adaptive fuzzy regression model for the prediction of dichotomous response variables, in: 15th Conference on Computational Science and Applications, Malaysia, 2007, pp. 14–19.

Ranjit K.p (2014), Some Methods of Detection of Outliers in Linear Regression Model, Iasri, *Library Avenue*, New Delhi-110012.

Redden D.T., Woodall W.H. (1994). Properties of certain fuzzy linear regression methods, *Fuzzy Sets and Systems, 64*, 361 – 375.

Salibian, M. and Yohai, V.J. (2006). A Fast Algoritm for S-Regression Estimates. *Journal of Computational and Graphical Statistics*,15, No. 2 (2006), 414-427.

Shiferaw, F., Letebo, M., Misganaw, A., Feleke, Y., Gelibo, T., Getachew, T. et al. (2018) Non communicable diseases in Ethiopia, disease burden, gaps in health care delivery and strategic directions. Ethiop J Health Dev. 32,3.

Stats, F. National chronic kidney disease fact sheet. (2017) US Department of Health and Human Services, Centers for Disease Control and Prevention.

Tanaka H. (1987). Fuzzy data analysis by possibilistic linear models. *Fuzzy Sets and Systems, 24*, 363375.

Tanaka H., Hayashi I., Watada J. (1989). Possibilistic linear regression analysis for fuzzy data. *European J. Oper. Res., 40*, 389-396. 44

Tanaka H., Uejima S., Asai K. (1982) Linear regression analysis with fuzzy model, *IEEE Transactions on Systems, Man and Cybernetics, 12*, 903–907.

Ulasi I. and Ijeoma, C, K. et al. (2010). highlighting the enormity of CKD in Nigeria with the situation in a Teaching Hospital in the southeast of Nigeria. Epub 2010(06) 2

Yager R. R. (1986) *An Introduction to Fuzzy Theory, Applications of Fuzzy Set Theory in Human Factors*, Elsevier Science, Amsterdam.