International STEM Journal, Volume 3 No.1, June 2022, 31-44

Development of a Fuzzy Model to Investigate HIV Threats

Ozoh Patrick[1*], Rasheed Sofiat[1], Ishola Esther[1] & Abanikannda Mutahir[2]

[1]Dept of Information and Communication Technology, Osun State University, Osogbo, Nigeria
[2]Dept of Science, Technology & Mathematics Education, Osun State University, Osogbo, Nigeria
[*]corresponding author: patrick.ozoh@umiosun.edu.ng

Received: 11 March 2022; Accepted: 27 June 2022

Abstract

This research develops HIV threat analysis model techniques for HIV risk analysis in a population for making informed decisions. The HIV disease is on the increase, and HIV/AIDS awareness may not be issued without actual threat investigation, to find out people's behavioral patterns. The motivation for this study is to find the solution to challenges in HIV infection. Hence, there is a need to develop a model that would investigate threats associated with the disease. This research investigates behavioral and demographic data to develop a fuzzy-based model. HIV reduces immunity against diseases that people with a strong immune system can ward off. This paper compares the results of the techniques tested on data obtained from the Virology Department of Obafemi Awolowo University Teaching Hospital. The techniques analyzed with MATLAB 2018b are adopted, and the triangular membership function was used for the system. The evaluation of the fuzzy model was achieved by comparing the technique with two past methods. This research indicates that the fuzzy logic technique will lead to getting the most accurate result.

Keywords: Fuzzy logic, HIV, Behavioral pattern, Threat analysis, Reliability.

Introduction

In recent data mining research, various algorithms have been applied to data sets from various sources to get information on HIV. Various techniques were used to establish relationships and discover new knowledge about HIV. The application of fuzzy logic to solve challenges in HIV infection is acknowledged because previous techniques have been applied to solve these problems. HIV is a medical challenge that is ravaging the world (Fauci & Lane, 2020). The previous techniques are included in this study.

The research papers have contributed immensely by using several approaches to predicting the likelihood of diseases. However, none has considered using behavioral and demographic data for their prediction. This research work focused on the design of a fuzzy logic-based model predictive solution to determine and measure the level of risk of HIV epidemics in a population using the information on their behavioral and demographic data. Hence, the contributions of this study are stated as follows:

1. Develops an HIV threat analysis model
2. Use of behavioral and demographic risk factors relating to HIV
3. Making informed decisions based on the developed model

The disease has contributed significantly to about 35.0 million deaths as of 2016 with sub-Sahara Africa as the worst with the greatest impact of this death (Amuche *et al*., 2017). There is currently no cure for the disease as it is transmitted through various means including sharing infected needles, contaminated blood transfusion, and practicing unsafe sex (Interim, 2005). The concept of fuzzy logic depends on human thinking and natural activities. The fuzzy logic works on the same level of possibilities of input, resulting in definite output (Oluwagbemi *et al*., (2020). Formation of a fuzzy logic knowledge-based known as the Fuzzy logic toolbox (Kovacic & Bogdan, 2018).

Recently, computational intelligence has been used on several complicated medical issues by developing intelligent systems (Adlassnig, 1986). Singla *et al.* (2020) did a study on kidney diseas diagnosis by using fuzzy logic. The paper developed a system in fuzzy logic for the diagnosis of kidney disease. Dagar *et al.* (2015) researched medical diagnosis using fuzzy logic and concluded that the technique of fuzzy logic can contribute to a reliable result to notify the disease. Annang (2018) predicted HIV status among female sex workers in Ghana using data mining techniques. The study, which used Random Tree, J48, Naïve Bayes, Logistic regression, and Neural Network algorithms determined if a female sex worker has ever tested for HIV. The paper adopted the CRISP-DM methodology and concluded that the best performance was achieved by a Random Tree classifier with an accuracy of 98.9%.

Adebayo *et al.* (2016) did a study on the prediction of pediatric HIV/AIDS patient survival. The paper demonstrated how the technique can be used to classify the survival of pediatric HIV/AIDS patients with an accuracy of 60% to 100% for the selected dependent variables. The research concluded that the naïve Bayes classifier was able to predict the survival of YES more accurately than the survival of NO and about 68% of the survival of NO was correctly classified. Young *et al.* (2017) did a study on how Machine learning approaches can be used to assist a domain expert who manually searches social data for HIV-related content. The paper used an existing social media dataset that was associated with HIV, four commonly used machine learning methods were tested to know whether they could learn the patterns associated with HIV risk behavior.

El-Halees and Shurrab (2017) worked on the prediction of Blood tumors using data mining techniques. The paper concluded that deep learning has the highest accuracy of prediction of blood tumors with an accuracy of 79.45%. Agarwal *et al.* (2019) researched the determination of risk factors for Hepatitis C. Isakki (2017) did a study on the prediction of HIV patients by a comparative study of data mining classification techniques using HIV/AIDS and STD data. The study used Decision Tree, Support Vector Machine (SVM), and Naïve Bayes' Classification algorithms. A study conducted by Mei and Zhao (2018) predicted HIV-1 and HIV-2 proteins using Chou's Pseudo amino acid compositions and different classifiers. Ozoh (2021) provides machines with the ability to collect data using systems, and then process collected data via machine learning methods for predictive purposes and inferring decisions at the same level as humans. Implementing accurate and reliable methods is important in modeling and predictions, otherwise poor results might occur. This research identifies the reliability of evolving predictive models utilizing mathematical models (Ozoh *et al.,* 2021).

The research papers have contributed immensely by using several approaches to predicting the likelihood of diseases. However, none has considered using behavioral and demographic data for their prediction. This research work identifies a fuzzy logic-based model predictive solution to determine and measure the level of risk of HIV epidemics in a population using the information on their behavioral and demographic data.

The fuzzy logic is the most accurate technique compared with the three previous methods considered in this study. The previous techniques considered are the regression method and the Box-Jenkins technique.

Methodology

The set of data was collected from the Virology Department of Obafemi Awolowo University Teaching Hospital. The first stage of data collection involves sourcing data from the virology department. The second stage will implement the techniques for evaluation. The predictive model was designed using FUZZY LOGIC and simulated in MATLAB/Simulink. The overall system design of the proposed study is shown in Figure 1.

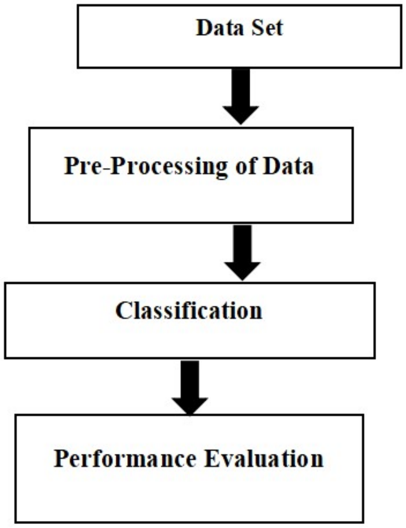


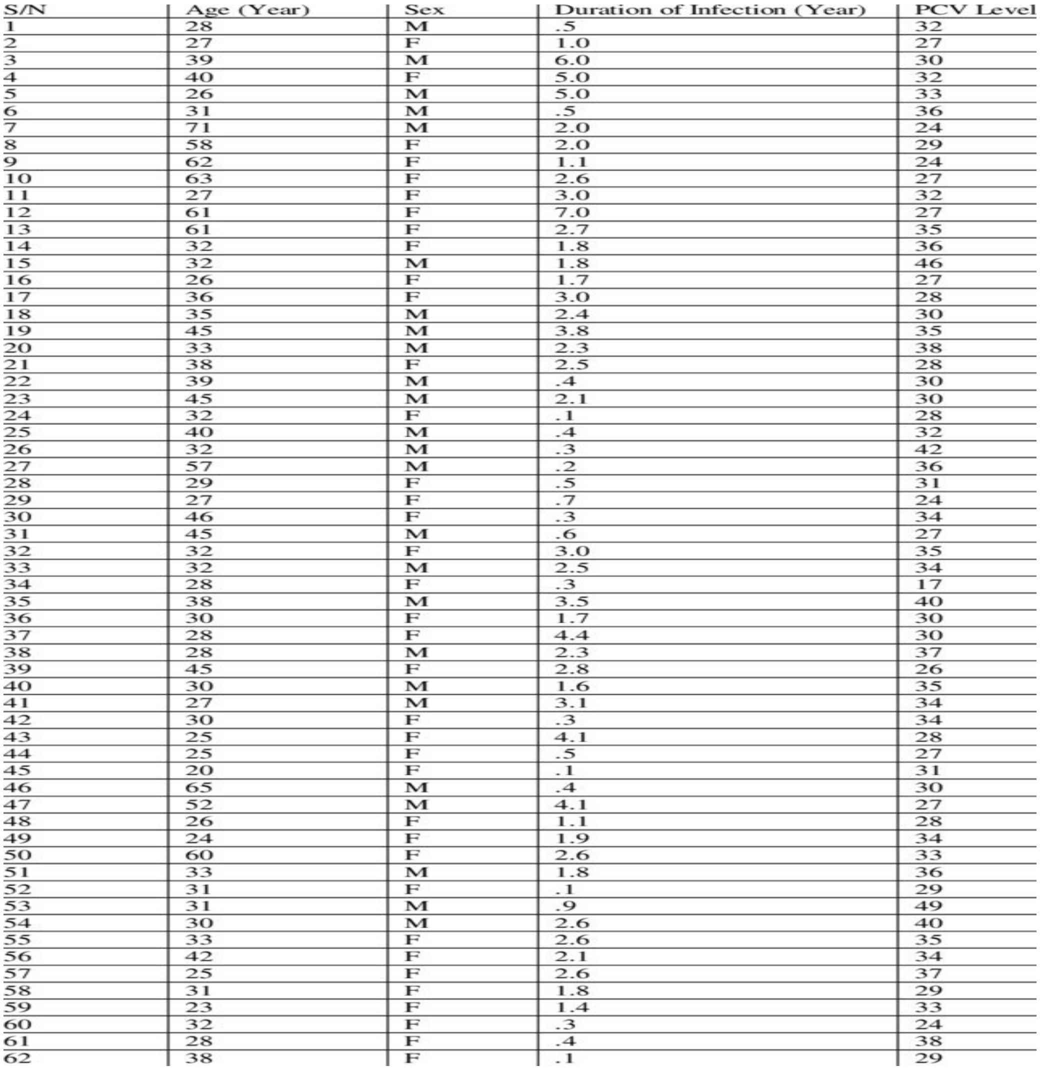
Figure 1: Proposed System Design

The evaluation will be carried out using behavioral and demographic data of the infected patient. This research models techniques for HIV risk prediction which will help individuals, in this case, patients performing consultations for self-analysis. This system will behave like an expert, like performing interviews with the patients. It offers many advantages to both medical practitioners and patients.

HIV Risk Dataset

This research has been carried out depending on knowledge gotten from experts from the Virology Department of Obafemi Awolowo University Teaching Hospital (OAUTHC) majorly on HIV. The dataset used for training is specified with a matrix with n+1 columns, where n represents input data into the fuzzy inference system and the final column indicates the output data for the system. The input data for HIV Risk used in this study are the symptoms which are 20 (twenty) attributes, some of which include chronic cough, headache, diarrhea, swollen gland, etc., with 1 (one) final column of the HIV risk.

A sample of 62 patients was taken showing their respective age, sex, duration of infection, and PCV level. The sample taken is given in Figure 2.



S/N	Age (Year)	Sex	Duration of Infection (Year)	PCV Level
1	28	M	.5	32
2	27	F	1.0	27
3	39	M	6.0	30
4	40	F	5.0	32
5	26	M	5.0	33
6	31	M	.5	36
7	71	M	2.0	24
8	58	F	2.0	29
9	62	F	1.1	24
10	63	F	2.6	27
11	27	F	3.0	32
12	61	F	7.0	27
13	61	F	2.7	35
14	32	F	1.8	36
15	32	M	1.8	46
16	26	F	1.7	27
17	36	F	3.0	28
18	35	M	2.4	30
19	45	M	3.8	35
20	33	M	2.3	38
21	38	F	2.5	28
22	39	M	.4	30
23	45	M	2.1	30
24	32	F	.1	28
25	40	M	.4	32
26	32	M	.3	42
27	57	M	.2	36
28	29	F	.5	31
29	27	F	.7	24
30	46	F	.3	34
31	45	M	.6	27
32	32	F	3.0	35
33	32	M	2.5	34
34	28	F	.3	17
35	38	M	3.5	40
36	30	F	1.7	30
37	28	F	4.4	30
38	28	M	2.3	37
39	45	F	2.8	26
40	30	M	1.6	35
41	27	M	3.1	34
42	30	F	.3	34
43	25	F	4.1	28
44	25	F	.5	27
45	20	F	.1	31
46	65	M	.4	30
47	52	M	4.1	27
48	26	F	1.1	28
49	24	F	1.9	34
50	60	F	2.6	33
51	33	M	1.8	36
52	31	F	.1	29
53	31	M	.9	49
54	30	M	2.6	40
55	33	F	2.6	35
56	42	F	2.1	34
57	25	F	2.6	37
58	31	F	1.8	29
59	23	F	1.4	33
60	32	F	.3	24
61	28	F	.4	38
62	38	F	.1	29

Figure 2: Sample Data of HIV Patients

Previous Techniques

The descriptions of past techniques used for developing models for HIV threats are explained in this section. These techniques include the regression method and the Box-Jenkins technique.

Box-Jenkins Technique

This technique used is based on Asamoah *et al.* (2012). it is used to predict future behavior. This is given as follows:.

$$\nabla zt = zt - zt - 1 = (1 - B)\ zt \qquad (1)$$

$$zt = \mu + at + \psi 1 at - 1 + \psi 2\ at - 2 + \ldots = \mu + \psi (B) at \qquad (2)$$

$\psi (B) = 1 + \psi 1\ B + \psi 2\ B2 + \ldots$

when $t - 1$, $t - 2, \ldots$ be $zt - 1$, $zt - 2$,...

$zt = zt - \mu$

Regression Technique

The regression technique used in this study is based on Wu and Yu (2018).

$$\varepsilon 2 = \Sigma\ (yi - yp)2 \qquad (3)$$

where yp = predicted values, and yi = actual values.

$$yp = m\ xi + b \qquad (4)$$

$$= 2m\Sigma xi2 + 2b\Sigma xi - 2\Sigma(xiyi) = 0 \qquad (5)$$

$$= 2m\Sigma xi + 2\Sigma b - 2\Sigma yi = 0 \qquad (6)$$

$Sxy = mSxx + bSx$ (12)

$Sy = mSx + nb$ (13)

HIV Risk Prediction System

The principle of this system has two major components which are symptoms as input and the output as the risk level. Figure 3 shows the HIV prediction system.

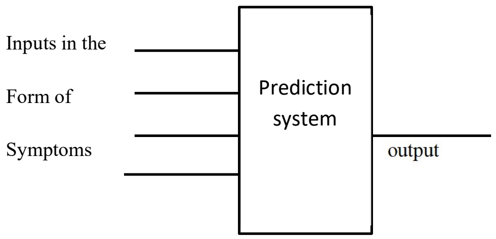


Figure 3: HIV Risk Prediction System

The Fuzzy logic system for developing models for HIV threats has four main parts, and is given as follows:

(1) Fuzzification module: It changes system variables.

(ii) Knowledge-based: It keeps the rules.
(iii) Inference engine: It performs inferences on the rules
(iv) Defuzzification module: It changes data obtained.

Figure 4 shows the architecture.

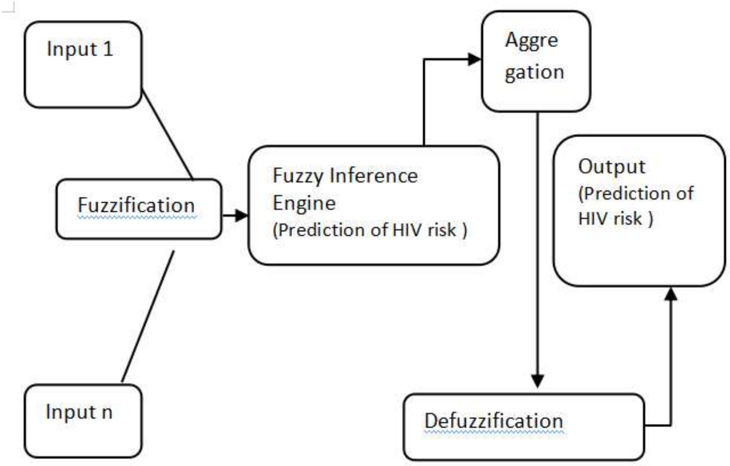


Figure 4: Fuzzy Logic Architecture

Consider that X signifies the space of discourse of a fuzzy set A. If fuzzy sets are selected, then a membership function for each set should be generated. The rules are distinctive curvatures that transfigure the input variables within a range signifying belongingness (Hale *et al.,* 2021). It can be represented by a collection of three points that form a triangle. Bräse *et al.* (2021) define membership function as a triangular membership function. The Triangular membership function was adopted in this research work. The triangular membership function can be expressed mathematically by Equation 7.

$$Tri(x;A,B,G) = \begin{cases} 0 & for\ x < A \\ \frac{x-A}{B-A} & for\ A \leq x \leq B \\ -\frac{x-G}{G-B} & for\ B \leq x < G \\ 0 & for\ x \geq G \end{cases}$$

(7)

It is considered that the output of each rule is also a fuzzy set. All the output fuzzy sets are then gathered and collected into one single fuzzy set. This process is called aggregation (Hale *et al,* 2021). In the end, the output set is determined to a single crisp number by defuzzification. Various defuzzification techniques are utilized for this study. It is done as shown by Equation 8.

$$x_{def} = \frac{\int \mu_A(x) x\, dx}{\int \mu_A(x)\, dx}$$

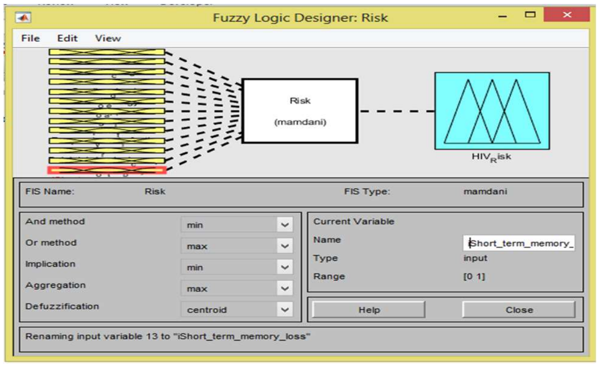
(8)

where x_{def} is the defuzzified output and μA (x)is the results after the aggregation of individual implication results. To develop fuzzy models, there is a need to have knowledge and understanding of the operations of the theory of fuzzy sets. The fuzzy set operations are defined concerning the sets' membership functions. This is shown by Equation 9.

□□□□X : μA(x) = μB(x)

$$\mu_A(x) \cap \mu_B(y) = \begin{cases} \min\{\mu_A(x), \mu_B(y)\} \\ \dfrac{\mu_A(x) + \mu_B(y)}{2} \\ \mu_A(x)\mu_B(y) \end{cases} \tag{9}$$

The FIS Toolbox is displayed in Figure 5.



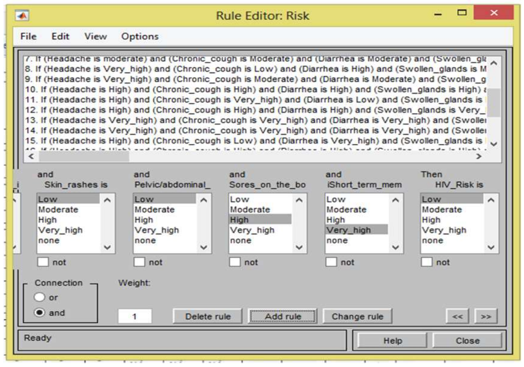
FIS Name: Risk
FIS Type: mamdani
And method: min
Or method: max
Implication: min
Aggregation: max
Defuzzification: centroid
Current Variable
Name: iShort_term_memory_
Type: input
Range: [0 1]
Renaming input variable 13 to "iShort_term_memory_loss"

Figure 5: Fuzzy Interface System Editor

For this research, 20 inputs were used for predictions, while 3 other attributes were used for the analyses which are gender, level of education, and age. Figure 6 shows the membership function of one of the inputs which is Headache. The range is set from 0 to 1, which means taking the values for the level at which a person can experience the symptoms.

Figure 6: Membership Function Editor for HIV Risk

Figure 7 shows the rule editor in MATLAB. The rules are constructed based on input and output.



```
7. If (Headache is moderate) and (Chronic_cough is Moderate) and (Diarrhea is Moderate) and (Swollen_gla
8. If (Headache is Very_high) and (Chronic_cough is Low) and (Diarrhea is High) and (Swollen_glands is M
9. If (Headache is Very_high) and (Chronic_cough is Moderate) and (Diarrhea is Moderate) and (Swollen_g
10. If (Headache is High) and (Chronic_cough is High) and (Diarrhea is High) and (Swollen_glands is High) a
11. If (Headache is High) and (Chronic_cough is Very_high) and (Diarrhea is Low) and (Swollen_glands is
12. If (Headache is High) and (Chronic_cough is High) and (Diarrhea is High) and (Swollen_glands is Very_
13. If (Headache is Very_high) and (Chronic_cough is Very_high) and (Diarrhea is Very_high) and (Swolle
14. If (Headache is Very_high) and (Chronic_cough is Very_high) and (Diarrhea is Very_high) and (Swolle
15. If (Headache is High) and (Chronic_cough is Low) and (Diarrhea is Very_high) and (Swollen_glands is
```

Figure 7: Rule Editor for HIV Risk

Figure 8 consists of deffuzification.

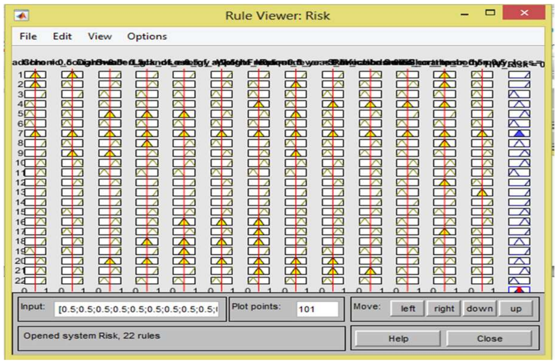


Figure 8: Rule Viewer for HIV Risk

Figure 9 shows the surface viewer which represents the concerning inputs

Figure 9: Surface Viewer for HIV Risk

Interpretation of Rules
For this prediction, 22 different rules were generated. Some of which include.

Rule1. If headache = chronic cough = moderate and diarrhea = very high and swollen gland = very high = loss of appetite = weight loss = very high and frequent fever = very high and frequent yeast infection = very high and skin rashes = high and pelvic/abdominal cramp = low and scores = moderate and short-term memory loss = low and Temperature =very high and Usage of sharp object = high and Drug usage =very high and Sex partner =very high and Transfusion of blood = Not sure and practice of safe sex = No and practice of circumcision = yes then HIV = very high

Rule2. If headache = moderate and chronic cough = low and diarrhea = high and swollen gland = high and lack of energy =low and loss of appetite = high weight loss = low and frequent fever = moderate and frequent yeast infection = high and skin rashes = low and pelvic/abdominal cramp = high and scores on certain parts of the body = moderate and short-term memory loss = low and Temperature =high and Usage of sharp object =very high and Drug usage =very high and Sex partner =high and Transfusion of blood = Not sure and practice of safe sex = No and practice of circumcision = yes then HIV = very high

Rule3. If headache = very high and chronic cough = very high and diarrhea = low and swollen gland = low and lack of energy = low and loss of appetite = low and weight loss = low and frequent fever = low and frequent yeast infection = low and skin rashes = low and pelvic/abdominal cramp = low and scores on certain parts of the body = low and short term memory loss = low and Temperature =low and Usage of sharp object = high and Drug usage =low and Sex partner =low and Transfusion of blood = screened and practice of safe sex = yes and practice of circumcision = yes then HIV = low

Rule4. If headache = high and Usage of sharp object =low drug usage =moderate sex partner =moderate and Transfusion of blood = unscreened and practice of safe sex = yes and practice of circumcision

Rule5. If headache = very high and scores on certain parts of the body = very high and short-term memory loss = very high. Temperature =high and Usage of sharp object =moderate. Drug usage =very high. Sex partner = high and Transfusion of blood = not sure and practice of safe sex = yes and practice of circumcision = yes then HIV = high

Results and Discussion

The accuracy of the system was analyzed on the prescription provided by experts and system-generated output. Varying symptoms of 56patients were provided to detect the exact possibility of distribution over the stage. The fuzzy basis technique was utilized, which potentially reduced the conservation of obtained result. The accuracy of the model was 91%. Our analysis was done by considering the age, gender, and level of education concerning the HIV risk. Fig 10 shows the bar chart of the HIV risk of the tested individuals.
Fig 11 shows the chart of the tested individual against the HIV risk

Fig 12 shows of the level of education of the tested individual against the HIV risk.

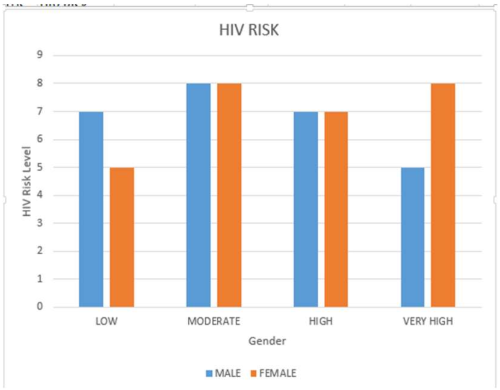


Figure 10: Chart for the Gender of the Tested Individual

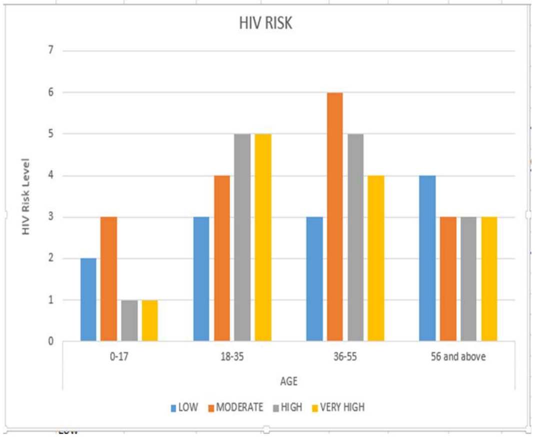


Figure 11: Chart for the Age of the Tested Individual

Figure 12: Chart for the Level of Education of Tested Individuals

Evaluation of Results

The actual data and estimates for sample data using the Box-Jenkins, Regression, and the fuzzy model are given in Table 1.

Table 1. Actual Data and Estimates for Techniques

Box-Jenkins		**Regression**		**Fuzzy**	
Actual data	**Estimates**	**Actual data**	**Estimates**	**Actual data**	**Estimates**
39.59	46.11	39.59	43.56	39.59	44.18
43.60	48.43	43.60	46.77	43.60	47.91
40.74	44.67	40.74	43.14	40.74	43.94
42.75	48.95	42.75	46.32	42.75	47.01
39.18	46.21	39.18	43.34	39.18	44.22
34.79	40.66	34.79	37.21	34.79	38.54
39.68	45.79	39.68	43.11	39.68	44.32
40.84	44.41	40.84	42.32	40.84	42.99
42.03	45.76	42.03	43.01	42.03	43.87
41.78	46.46	41.78	43.32	41.78	44.58
42.94	46.42	42.94	43.11	42.94	44.36
38.12	43.68	38.12	39.92	38.12	40.93
35.91	40.43	35.91	37.88	35.91	38.03
40.95	45.74	40.95	42.43	40.95	43.46
41.56	47.32	41.56	44.47	41.56	45.84
42.34	45.46	42.34	43.13	42.34	43.79
42.64	46.34	42.64	43.73	42.64	44.74
42.15	46.85	42.15	43.58	42.15	44.36
36.90	40.85	36.90	37.41	36.90	38.58
35.76	39.47	35.76	36.74	35.76	37.74
41.58	45.27	41.58	42.82	41.58	43.43
42.15	46.11	42.15	43.15	42.15	44.85
40.92	44.32	40.92	41.89		

The results of the comparisons are shown in Table 2.

Table 2. Evaluation of Techniques

	Box-Jenkins	**Fuzzy**	**Regression**
RMSE	0.873	0.596	0.704
MAPE (%)	1.957	0.921	1.199

The output for the RMSE and MAPE of Box-Jenkins, fuzzy, and the regression method indicate that the fuzzy model is the most accurate for estimating data. The RMSE and MAPE values of the fuzzy model are the smallest evaluated with Box-Jenkins and regression. Table 2 indicates that RMSE for Box-

Jenkins, fuzzy model and regression are 0.873, 0.596, and 0.704 respectively. The MAPE for the fuzzy model is 0.921%, which is the smallest compared with Box-Jenkins, and regression with MAPE values of 1.957%, and 1.199% respectively.

Conclusion

The risk of having HIV in this modern time is on the increase and HIV/AIDS sensitization may not be fully provided without proper risk analysis prediction or behavioral pattern check of people in a particular region. This study, therefore, aimed at designing an HIV risk prediction model with intention of providing tools for HIV risk analysis in a population for making informed decisions. The proposed system was designed using behavioral and demographic risk factors relating to HIV outbreaks. Fourteen symptoms of HIV risk are Headache, Chronic cough, Diarrhea, Swollen gland, Lack of energy, loss of appetite, Weight loss, Frequent fever, Frequent yeast infection, Skin rashes, pelvic/abdominal cramps, Sores on certain parts of the body, short term memory and Temperature. Also, ten behavioral data were used which include Blood transfusion, Practice of safe sex, Practices of circumcision, Usage of sharp objects, Drug usage, Level of Education, Sex partner, Age, and Marital status.

These attributes were categorized based on severity level as Low, Moderate, High, and Very high which serve as input to the fuzzy inference system. A fuzzy logic toolbox in MATLAB 2018b was adopted for designing and modeling the system. A triangular membership function was used. The proposed system will be a useful tool for individuals, government, researchers, and health workers in predicting and analyzing the HIV risk of a population in a particular area.

References

Adebayo, I. P., Olutola, A. and Aladekomo, T. A. (2016). The Prediction of Paediatric HIV/AIDS Patient Survival: A Data Mining Approach. *Asian Journal of Computer and Information Systems*, 4(3).

Adlassnig, K. P. (1986). Fuzzy set theory in medical diagnosis. IEEE *Transactions on Systems, Man, and Cybernetics*, 16(2), 260-265.

Agarwal, G. G., Singh, A. K., Venkatesh, V. and Wal, N. (2019). Determination of risk factors for hepatitis C by the method of random forest. *Annal of Infectious Disease and Epidemiology*, 4(1).

Amuche, N. J., Emmanuel, E. I. and Innocent, N. E. (2017). HIV/AIDS in sub-Saharan Africa: current status, challenges and prospects.

Annang, D. A. (2018). Performance Comparison of Data Mining Techniques for Predicting Hiv Status Among Female Sex Workers in Ghana (Doctoral dissertation, University of Ghana).

Asamoah, D., Annan, J. and Arthur, Y. D. (2012). Time Series Analysis of Electricity Meter Supply in Ghana. *International Journal of Business and Social Science*, 3(19).

Bräse, S., Gil, C., Knepper, K. and Zimmermann, V. (2005). Organic azides: an exploding diversity of a unique class of compounds. *Angewandte Chemie International Edition*, 44(33), 5188-5240.

Dagar, P., Jatain, A. and Gaur, D. (2015). Medical diagnosis system using fuzzy logic toolbox. In International Conference on Computing, Communication & Automation (pp. 193-197). IEEE.

El-Halees, A. M. and Shurrab, A. H. (2017). Blood tumor prediction using data mining techniques. *Health Informatics—An International Journal*, 6.

Fauci, A. S. and Lane, H. C. (2020). Four decades of HIV/AIDS—much accomplished, much to do. *New England Journal of Medicine*, 383(1), 1-4.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T. and Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4), 529-538.

Interim, W. H. O. (2005). clinical staging of HIV/AIDS and HIV/AIDS case definitions for surveillance: African region. World Health Organization.

Isakki, S. L. D. (2017). A Comparative Study of Data Mining Classification Technique Using HIV/AIDS and STD data. *Int. J. Innov. Res. Comput. Commun. Eng*, 5(1), 134-139.

Kovacic, Z. and Bogdan, S. (2018). Fuzzy controller design: theory and applications. CRC press.

Mei, J. and Zhao, J. (2018). Prediction of HIV-1 and HIV-2 proteins by using Chou’s pseudo amino acid compositions and different classifiers. *Scientific Reports,* 8(1), 1-9.

Singla, J., Kaur, B., Prashar, D., Jha, S., Joshi, G. P., Park, K. and Seo, C. (2020). A Novel Fuzzy Logic-Based Medical Expert System for Diagnosis of Chronic Kidney Disease. *Mobile Information Systems,* 2020.

Oluwagbemi, O. O., Oluwagbemi, F. E., Jatto, A. and Hui, C. (2020). MAVSCOT: A fuzzy logic-based HIV diagnostic system with indigenous multi-lingual interfaces for rural Africa. *PloS one*, 15(11), e0241864.

Ozoh, P. (2021). A comparison of machine learning techniques via computational methods. *Annals. Computer Science Series*, 19(1).

Ozoh, P., Olayiwola, M. and Adegoke, A. (2021). Appraisal of Predictive Techniques using Computational Methods. *International STEM Journal,* 2(2).

Wu, C. and Yu, J. Z. (2018). Evaluation of linear regression techniques for atmospheric applications: the importance of appropriate weighting. *Atmospheric Measurement Techniques*, 11(2), 1233-1250.

Young, S. D., Yu, W. and Wang, W. (2017). Toward automating HIV identification: machine learning for rapid identification of HIV-related social media data*. Journal of acquired immune deficiency syndromes* (1999), 74(Suppl 2), S128.