Insights into gender-equity in healthcare accessibility in Northern Nigeria: descriptive and predictive approaches

Chika Yinka-Banjo,[1] Olasupo Ajayi,[2] Mary Akinyemi,[3] David Tresner-Kirsch,[4,5] Adekemi Omotubora[6]

[1]Department of Computer Sciences, University of Lagos, Lagos, Nigeria; [2]CAESAR Lab, Queen's University, Kingston, ON, Canada; [3]Department of Mathematics & Statistics, Austin Peay State University, Clarksville, TN, United States; [4]Nivi, Inc., Sudbury, MA, United States; [5]Department of Computer Science, Brandeis University, Waltham, MA, United States; [6]Department of Commercial & Industrial Law, University of Lagos, Lagos, Nigeria

Abstract

Universal Health Coverage (UHC) aims at ensuring equitable access to healthcare for everyone, irrespective of gender, location, or financial status. Though progress has been made in achieving UHC, a lot remains to be done in under-served areas of the world. These regions face immense challenges accessing healthcare services, including unavailability of basic medications, socio-cultural and religious beliefs, and various forms of discrimination. Beyond these, women are still severely disadvantaged in these regions, with child brides and teenage pregnancy being prevalent. This work analysed data from regions of northern Nigeria to determine equity in healthcare accessibility. Descriptive analysis (using correlation models) and predictive analysis (using machine learning models) were carried out. The descriptive analysis revealed that women with low income and education levels, and the elderly have a higher chance of accessing healthcare services compared to other genders, while the predictive analysis revealed that, using machine learning, accessibility to healthcare services can be predicted with up to 81% accuracy.

Introduction

In the last few years, there has been an overarching improvement in healthcare services globally. Reports from the United Nation and World Health Organisation show that significant efforts have been made in reducing deaths resulting from various diseases, as well as infant and maternal mortality globally.[1,2] Despite these efforts, healthcare accessibility remains a challenge in many parts of the world, specifically the global south countries of Africa and Asia. These regions have the lowest number of healthcare professionals at 2.3 medical doctors per 10,000 people compared to the world average of 16.3 doctors per 10,000 people. Women are the most affected by the poor healthcare services in these regions. For instance, approximately 90% of all maternal mortalities globally were in sub-Saharan Africa, Central and Southern Asia. Similarly, these regions also have the highest percentage of adolescent pregnancy (*i.e.*, pregnant girls between the ages of 15 and 19).[1]

Several studies have attempted to assess equity in healthcare systems across Africa. In some work, statistical models were used,[3-5] while others used Artificial Intelligence (AI) and Machine Learning (ML) models to analyse health data,[6] predict the outcome of healthcare services,[7] assess gender equity,[8] and/or eliminate gender biases in healthcare.[9-11] Like these works, this study also seeks to assess gender-equity in healthcare but from both descriptive and predictive perspectives, considering individual

Correspondence: Olasupo Ajayi, CAESAR Lab, Department of Electrical and Computer Engineering, Queen's University, Kingston, K7L 3N6, Ontario, Canada; +1437-421-8617, o.ajayi@queensu.ca

Key words: data analysis, data correlation, gender-equity, healthcare, machine learning.

Contributions: CY-B, principal investigator, conceptualization, methodology; OA, investigation, software, writing – original draft; MA, co-investigator, formal analysis, and validation; DT-K, resources, data collection, administration; AO, writing reviewing & editing, adherence to ethics.

Ethics approval and consent to participate: College of Medicine, University of Lagos - Health Research Ethics Committee. CMUL/HREC/11/22/1130. Data were collected from respondents who gave their consent to participate. All respondents were presented with a first page on the Nivi app, which clearly stated that the data being collected was for research and analytic purposes and would remain completely anonymous without any form of tracking. All respondents were required to accept these terms and conditions before being allowed to participate in the data collection exercise. Data were collected using askNivi, an app specifically designed for healthcare related data collection. Privacy policy, consent, and related information about askNivi can be found at www.nivi.io/privacy-policy

Availability of data and material: the datasets generated and/or analysed during this study can be found in the Mendeley repository, https://data.mendeley.com/datasets/8gbywtd7bv/2.

Conflict of interest: the authors declare that they have no competing interests.

Funding: United States Agency for International Development (USAID) and the DAI Global, under the 'Nivi and the University of Lagos (UNILAG): Partnering to Create a Gender-Aware Auditing Tool' project.

Acknowledgement: the authors would like to acknowledge Nivi Inc. for the use of their askNivi application platform to administer the questionnaires, as well as the Artificial Intelligence and Robotics Lab (AiRoL) at the University of Lagos, for the use of their space for collaborative meetings.

Received: 28 November 2024.
Accepted: 23 May 2025.
Early access: 9 July 2025.

Healthcare in Low-resource Settings 2025; 13:13433
doi:10.4081/hls.2025.13433

characteristics (age, gender), socio-economic factors (income, education, and marital status) and medical symptoms as input parameters. The specific contributions of this work are: i) to carry out descriptive analyses and assess the level of gender-equity in healthcare accessibility in northern Nigeria; this is especially relevant to goals 3 (good health and well-being) and 5 (gender equality) of the UN's Sustainable Development Goals (SDG); ii) to develop models for predicting healthcare accessibility which consider individual characteristics and socio-economic factors. These models seek to predict if an individual A, who is of gender G, marital status M, education level E, and income level I, would be able to access health care service.

Materials and Methods

In assessing healthcare accessibility, we developed the Gender-Equity Healthcare Access Framework (GEHAF) - a multi-phased framework depicted in Figure 1. Though the primary focus of this work is on Phase 3c, 4 and 5c of the framework, all the 5 phases are briefly described in this section. Three datasets were used for this work, however, for brevity, only the results of the first dataset are presented.

In phase 1 of Figure 1, data on gender-equitable accessibility to healthcare was collected as follows: i) data for this work was collected from about 500 towns and villages across the northern region of Nigeria. The collection area spanned geographical coordinates: 13.032880, 5.365820 and 9.267890, 5.345540 to the west to 11.831740, 13.166820 and 9.173710, 12.415130 to the east; ii) data was collected using questionnaires administered via a mobile app called askNivi.[18] iii) askNivi is a conversational chat-like health tool developed to provide health education and referrals for healthcare systems; iv) all respondents gave their consent before participating in the data collection exercise; v) the respondents were asked 43 questions, split into four sections, a) demographic information, b) socio-economic information, c) wellness check, d) healthcare DEI. The questions were a mix of Boolean ('Yes/No'), short answer / free text, and Likert scale ('agree', 'neutral', 'disagree') type questions; and were designed to be simple and unambiguous; vi) the 43 questions corresponded to the 43 attributes in the dataset, namely: 'age', 'gender', 'language', 'askmarital', 'asklocation', 'askhavekids', 'howmanykids', 'nationality', 'ethnicity', 'askeducation', 'askincome', 'askemployment', 'healthaccess', 'typeofhealthcare', 'socialsupport', 'details', 'askfamilysupport', 'familysymptoms', 'wellness', 'symptoms', 'howlongsymptoms', 'gettingbetterorworse', 'whichfamily', 'sex', 'partners', 'overallhealth', 'medicines', 'medicinescurrent', 'medicinesrecent', 'alcohol', 'smoke', 'howoftensmoke', 'howoftenalcohol', 'surgeries', 'describesurgeries', 'inclusion', 'equity', 'diversity', 'genderjusticesafety', 'harassment', 'genderjusticecomfort', 'discrimination', 'refusedtreatment'; vii) a total of 4,628 records across the 43 attributes were collected. The corresponding datasets, code book, and questionnaire are available in.[12]

The collected data was then pre-processed to remove duplicates, nulls, or invalid entries. Phase 1 also has a language translation module that helps translate data in any source language to English language.

In phase 2 and 3a, the 'symptoms' field - one of the 43 attributes in the dataset, was passed through Scispacy Named Entity Recognition (NER)[13] to extract medical related terms. The output of these phases was an annotation table of symptoms, as shown on Table 1.

In Phase 3b, a human auditing dashboard[14] was developed to improve the annotation accuracy in phase 3a. The revised annotation table was appended to the original dataset and used to train ML models in phase 3c.

In Phase 3c, the aim was to develop ML models capable of predicting healthcare accessibility. Specifically, it sought to determine if an individual of a certain gender and socio-economic class, who presents with certain symptoms (S), at a given healthcare facility in northern Nigeria would be treated or not. We formally model this as the quintuple in (1):

$$P = A\langle G, M, E, I, S_x \rangle \tag{1}$$

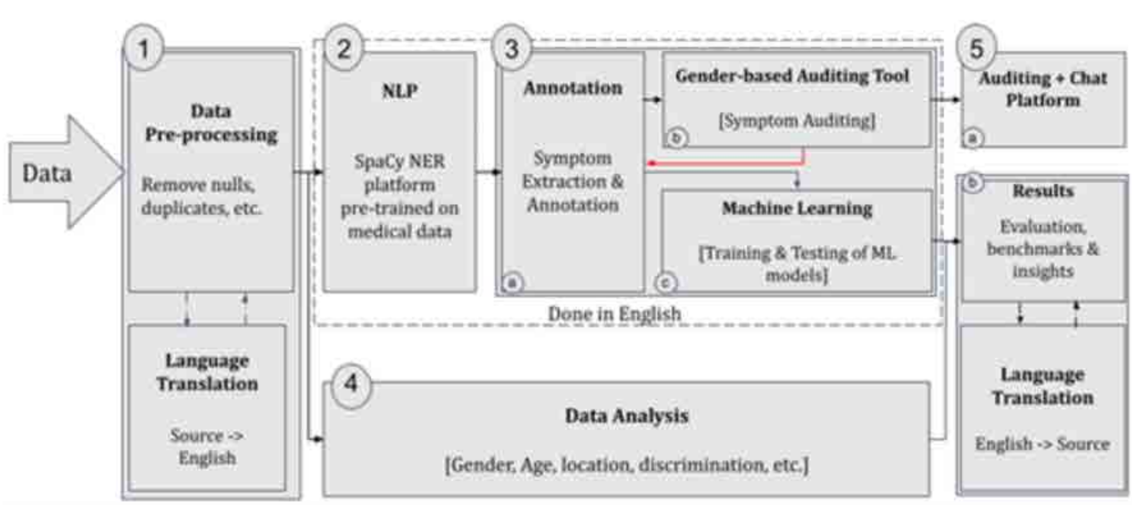

Figure 1. Gender Equity Healthcare Access Framework (GEHAF).

where A is an individual, G and M are A's gender and marital status respectively, while E is the level of education (ranging from primary to tertiary). I is A's income level, while S is a set of x symptoms, $x = \{1, 2 \ldots 17\}$, $\forall\ S_x = \{0, 1\}$, with 1 implying that the individual A is experiencing a particular symptom x, and 0 if otherwise. $P = \{0, 1\}$ and represents the prediction result, i.e., 0 if A would be denied health care or 1 if A would be able to access healthcare service.

To develop the prediction models, 6 of the 43 attributes in the dataset were considered, namely 'gender', 'askmarital', 'askeducation', 'askincome', 'symptoms', and 'refusedtreatment'. The following steps were then taken. The respondents' genders (man, woman, or non-binary) were extracted from the dataset. The gender attribute was label encoded into categorical data.

Similar or related symptoms were combined. For instance, symptoms such as 'dizziness' and 'dizzyness', were combined, as well as 'back pain' and 'backpain'; 'headache', 'head ache', and 'headaches'; 'itches', 'itchy', and 'itching', etc.

The most common (highest respondent count) symptoms in the dataset were selected. These were 'Stomach', 'Back', 'Chest', 'Coughs', 'Diabetes', 'Dizziness', 'Fever', 'Headaches', 'Body pains', 'Vomiting', 'Itches', 'Malaria', 'Teeth', 'Ulcer', 'Virginal', 'Catarrh', and 'Infections'.

Three socio-economic features were extracted from the dataset, namely 'askmarital' (marital status), 'askeducation' (level of education), and 'askincome' (level of income).

The socio-economic features were converted into categorical data using One-hot encoding.

Each respondent could choose 1 of 5 options when responding to the question about accessibility to healthcare (i.e., the 'refusedTreatment' attribute). These options were - 'Agree', 'Moderately Agree', 'Neutral', 'Moderately Disagree', and 'Disagree'. While processing this field 'Agree' and 'Moderately Agree' were combined into 'Agree', 'Moderately Disagree' and 'Disagree' were grouped as 'Disagree', while Neutral responses were dropped. This reduced the respondents to two groups. 'Agree' meant that the respondent was refused health care at the healthcare facility, while 'Disagree' meant that the respondent received health care. The 'refusedTreatment' attribute was the target column for the ML models and was label encoded into a binary class (1 or 0).

Three ML models were trained on the data and their performances compared in terms of prediction accuracy, precision, recall, and F1-score. Details of the ML models are presented in subsection "*Machine Learning Model*".

The final dataset used in training the ML models consisted of 22 columns summarised on Table 2.

In phase 4, the datasets were analysed using statistical models to provide insights on gender-equality in healthcare across regions of northern Nigeria. These descriptive analyses were done in parallel to phases 2 and 3 and are presented in subsection "*Data analysis*" Phase 5 presents the obtained results.

Data analysis

To analyse the relationship between individual characteristics (age, gender), socio-economic factors (marital status, location, level of education, income level, employment status, and ethnicity) and healthcare equity, (inclusion, equity, diversity, gender justice, harassment, discrimination, and access to treatment), statistical correlation analyses were carried out. We used the Pearson Correlation Coefficient (PCC) and Spearman Correlation Coefficient (SCC),[16] while the probability value (p-value) was used to measure the level of statistical importance.

Pearson Correlation Coefficient (PCC), designated as r, is

Table 1. Sample output of Phases 2 and 3a of GEHAF Framework.

Input (Symptoms)	Output of Phase 3a (Symptom Annotation Table)				
		Headache	Stomach Ache	Stooling	Fever
'I constantly have headaches and stomach aches in the mornings'	Record 1	1	1	0	0
'Stomach ache, stooling and fever since last week'	Record 2	0	1	1	1

Table 2. Dataset attributes.

SN	Attributes	Options
1	'Gender' (G)	Man, Non-binary, Woman
2	'askmarital' (M)	Divorced, Married, Single, Widow/Widower
3	'askeducation' (E)	Primary, Secondary, Tertiary, or Other
4	'askincome' (I)	Level 1 (>USD 100,000) - Level 6 (< USD 500)
5-21	'Stomach', 'Back', 'Chest', 'Teeth', 'Coughs', 'Diabetes', 'Dizziness', 'Fever', 'Ulcer', 'Headaches', 'Body pains', 'Vomiting', 'Itches', 'Malaria', 'Virginal', 'Catarrh', 'Infections'	1 or 0. 1 if the respondent experiences the symptom or 0 if otherwise.
22	'refusedTreatment'	Agree or Disagree

often used to find the association between two parametric attributes (e.g. age). PCC is calculated using (2) for attributes X and Y .

$$r = \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \right] \qquad (2)$$

where r = PCC, x' and y' are sample mean values of x and y respectively.

Spearman ranks Correlation Coefficient (SCC), denoted R_s, is used to show association between two non-parametric attributes, e.g. Likert scale. SCC is defined in (3).

$$R_s = 1 - \left[\frac{6 \sum d^2}{(n^3 - n)}\right] \qquad (3)$$

where *d* is the difference in rank between values of the two attributes being considered.

For PCC and SCC, the results usually fall within the range of $-1 \leq cc\_value \leq 1$. A *cc_value* between 0 and 0.39 (or between 0 and -0.39) implies a weak positive (negative) relationship between the two attributes, a *cc_value* between 0.4 and 0.69 (or between -0.4 and -0.69) implies a moderate positive (negative) relationship, while values between 0.7 and 1 (or between -0.7 and -1) imply strong positive (negative) relationship between the two attributes. Finally, we calculated the probability value (p-value) to determine the statistical significance of the correlation between two attributes and set the threshold to 0.05.

Machine learning models

The 'refusedTreatment' attribute was used as the target for the ML models. By grouping all values of the 'refusedTreatment' attribute into either 'Agree' and 'Disagree', the problem was reduced to a binary classification.

This dataset initially contained 1093 records, however, after the cleaning, preprocessing and annotation processes (phases 1-3 of GEHAF), only 412 valid records were left. The data was also highly imbalanced, with the majority (381) of records belonging to the 'refusedTreatment' = 'Disagree' class, *i.e.*, those who were able to access health care, while only a minority (31) of data belonged to the 'Agree' class (those who were refused treatment). To balance the dataset, we oversampled the minority class.[16]

Three ML models were considered in this work - CATBoost, Random Forest, and Support Vector Classifier (SVC). CATBoost or Categorical Boosting, is an unbiased gradient boosting algorithm that is suited for categorical data.[17] CATBoost was selected because our dataset was small and consisted of attributes with a mix of data types (Boolean, texts, and Likert values). Random Forest (RF) is based on multiple (ensembled) Decision Trees (DT). RF considers the result of each DT, then uses a voting system to make the final decision. Like DT, RF does not require prior information about the distribution of the data,[7] hence, suitable for our work. Support Vector Classifier (SVC) creates hyperplane(s) to separate variables into classes. We considered SVC for this work because it has good learning and generalisation abilities with small and heterogeneous datasets.

Two sets of experiments were carried out. For the first, the raw imbalance data was used, while the augmented (oversampled) dataset was used in the second. The experiments were done in Google Colab, using a combination of Python, Jupyter notebook, and SciKit-Learn library.[18] All codes used in this study are available in Code Ocean code repository.[19] Accuracy, precision, recall, and F1-score, were used as metrics. Recall and F1-score are best suited for measuring imbalance datasets, as Recall reveals the correctly detected proportion of true positives, while F1-Score is the harmonic mean of precision and recall.

Parameter tuning

Cross validation and hyper-parameter tuning for each of the ML models were done using SciKit-Learn's StratifiedKFold library.[18] For CATBoost and RF, we considered 4 iteration values - 250, 500, 1000, 2000; 3 depth values - 6, 8, 10; and 5 folds. F1-score was used as the evaluation metric for the cross-validation process. In addition, we also considered 3 learning rates for CATBoost, i.e., 0.1, 0.05, and 0.01. For SVC, the same iteration values were considered, alongside 3 kernel options - 'linear', 'rbf', and 'sigmoid'; 3 C-values - 0.01, 0.001, and 0.0001; 4 gamma values - 0.1, 1, 10, and 100. Table 3 summarises the combinations of parameters that yielded the best results and used in this work.

Results

As previously mentioned, two types of analyses were carried out on the dataset - descriptive data analysis using mathematical and statistical models, and predictive data analytics, using ML models. Reports of both are presented in this section.

Associations

All attributes in the dataset, except age, were one-hot encoded. For instance, with the 'gender' attribute, women were encoded as 0, non-binary as 1 and men as 2. Similar encoding scheme was used for all the other attributes. The associations between the attributes are shown on Table 4.

Table 3. Machine Learning parameter tuning.

	Raw Data	Augmented Data
CATBoost	Iterations = 250, Depth = 6, Learning rate = 0.01. F1-Score = 95.8 %	Iterations = 500, Depth = 10, Learning rate = 0.1. F1-Score = 77.7 %
SVC	Tolerance = 1e-05, Max_iteration = 250, gamma = 0.1, C = 0.01, kernel = 'rbf'. F1-Score = 95.8 %	Tolerance = 1e-05, Max_iteration = 500, gamma = 10, C = 0.01, kernel = 'rbf'. F1-Score = 79.8 %
RF	Iteration = 250, max_depth = 10. F1-Score = 95.2 %	Iteration = 500, max_depth = 10. F1-Score = 76.8 %

Descriptive analysis

The impact of gender and socio-economic attributes on healthcare accessibility are presented in this section.

Influence of age and gender on healthcare accessibility

Figure 2a shows that most (94%) of the respondents were below 40 years old. Across all the age groups, the majority (63% of those below 21, 61% of those between 21 and 30, 70% of those between 30 and 41 years old) disagreed with being refused treatment, i.e. they were able to access treatment from the healthcare facility.

In Figure 2b, 245 of the 318 women (77%) were able to access healthcare services, while 7% had challenges accessing healthcare services. Similarly, 68% of the men were able to access healthcare services, compared to 5% who could not. Only 2 of the 9 non-binary respondents had no challenges accessing health services. Overall, most (77% of the women, 68% of the men, 78% of the non-binary gender) of the respondents were able to access health services, a fact buttressed on Table 4 with a p-value of 0.016 indicating a statistically insignificant association between gender and healthcare accessibility.

Influence of location on healthcare accessibility

In Figure 3, 19 of the 277 (6.8%) women living in towns were refused treatment, compared to 4 of the 41 (9.8%) living in villages. Similarly, 3.8% of the city-dwelling men were refused healthcare services compared to 12.5% of the village dwellers. This shows that a higher percentage of village dwellers were unable to access healthcare services, compared to their town-dwelling counterparts.

Influence of marital status on healthcare accessibility

In Figure 4, 5.8% of the single women were refused treatment, compared to 10.3% of married women. For the men, 5.4% of the single men and 4.7% of the married men were unable to access treatment. Marriage is therefore not a critical requirement to accessing healthcare. With an associative value of -0.053 and a p-value of 0.249, Table 4 also shows that Marital Status and access to treatment ("refusedTreatment") are weakly associated.

Influence of income on gender-based healthcare accessibility

In Figure 5, level 1 denotes respondents with annual income

Table 4. Associations among attributes in the dataset.

	Age	Gndr	Mar	Loc	Edu	Emp	Incm	Ethn	Incl	Equi	Dvr	Gndr Just.	Hrsm	Dscr	RfTmt
Age	1	-0.226 (<.001)	-0.466 (<.001)	0.149 (<.001)	-0.017 (0.711)	-0.057 (0.215)	-0.180 (<.001)	0.006 (0.903)	0.056 (0.218)	-0.005 (0.911)	0.084 (0.064)	0.048 (0.289)	0.035 (0.442)	0.046 (0.313)	0.044 (0.340)
Gndr		1	-0.025 (0.573)	-0.035 (0.450)	0.027 (0.552)	0.130 (0.004)	0.064 (0.158)	0.109 (0.016)	-0.086 (0.059)	-0.040 (0.377)	-0.085 (0.062)	-0.078 (0.089)	-0.082 (0.072)	-0.076 (0.094)	-0.109 (0.016)
Mar			1	-0.052 (0.253)	-0.023 (0.613)	-0.001 (0.980)	0.137 (0.003)	-0.067 (0.140)	-0.046 (0.318)	-0.003 (0.946)	-0.064 (0.162)	0.017 (0.708)	-0.048 (0.292)	-0.056 (0.221)	-0.053 (0.249)
Loc				1	0.006 (0.892)	0.038 (0.411)	0.093 (0.041)	-0.096 (0.035)	0.042 (0.363)	0.009 (0.828)	0.107 (0.019)	0.107 (0.019)	0.163 (<.001)	0.129 (0.005)	0.152 (0.001)
Edu					1	0.081 (0.076)	0.008 (0.867)	-0.026 (0.571)	0.056 (0.223)	0.061 (0.184)	-0.001 (0.989)	0.0156 (0.733)	-0.038 (0.399)	-0.105 (0.022)	-0.096 (0.034)
Emp						1	0.049 (0.278)	0.065 (0.154)	-0.0662 (0.140)	-0.035 (0.445)	-0.020 (0.658)	-0.069 (0.127)	-0.058 (0.204)	-0.082 (0.074)	-0.020 (0.656)
Incm							1	-0.023 (0.609)	-0.005 (0.909)	0.049 (0.281)	0.090 (0.047)	-0.063 (0.169)	-0.038 (0.404)	0.029 (0.532)	0.089 (0.049)
Ethn								1	0.032 (0.482)	0.053 (0.246)	0.016 (0.721)	-0.029 (0.528)	-0.076 (0.095)	-0.044 (0.335)	-0.122 (0.007)
Incl									1	0.439 (<.001)	0.285 (<.001)	0.275 (<.001)	0.229 (<.001)	0.184 (<.001)	0.179 (<.001)
Equ										1	0.479 (<.001)	0.462 (<.001)	0.353 (<.001)	0.249 (<.001)	0.245 (<.001)
Dvr											1	0.595 (<.001)	0.375 (<.001)	0.419 (<.001)	0.366 (<.001)
Gndr Just.												1	0.424 (<.001)	0.499 (<.001)	0.452 (<.001)
Hrsm													1	0.569 (<.001)	0.464 (<.001)
Dscr														1	0.626 (<.001)
RfTmt															1

greater than NGN 70 million (USD 88,000). Level 2 are those earning between NGN 45 - 70 million or (USD 56,000 - 88,000), level 3 earned between NGN 25 - 44 million (USD 31,000 - 55,000), level 4 earned between NGN 7 - 24 million (USD 8,750 - 30,000), level 5 earned between NGN 0.7 - 6 million (USD 875 - 7,500), while level 6 implies an income of less than NGN 300,000 (USD 375). Most of the respondents were within levels 5 and 6. This is unsurprising as the minimum wage in Nigeria, as at the time of writing, was about NGN 360,000 (USD 450) annually. Due to the limited data about higher income levels (level 1 - level 4), only responses from levels 5 and 6 respondents were analysed.

Of the 245 level 6 women respondents, 16 (6.5%) were refused treatments, while 8 (11%) of the 55 level 5 women were refused treatment. For both income levels, over 78% were able to access treatment. For the men, 7% of those in level 6 were denied treatment, while none of the level 5 respondents were refused treatment. Of the 8 non-binary respondents, only 2 reported being refused treatment. Thus, for men, income level influences healthcare access, as the richer men (level 5 and 6) were almost never refused treatment. For the women, the reverse is observed, as more level 5 women were denied treatment compared to level 6 women. Statistically, an SCC value of 0.089, shown on Table 4, indicates a weak association between Income and “refusedTreatment”. In addition, the low p-value (0.049) implies that wealth might not always guarantee access to healthcare, as indicative of the reverse trend observed with more richer women (level 6) being denied treatment than any other income group.

Influence of employment on gender-based healthcare accessibility

Figure 6 reveals that only 6 of the 116 unemployed women were refused treatment, compared to the 83% who were able to access healthcare services. Similar ratios were observed for the employed women. Self-employed women had the most challenge accessing

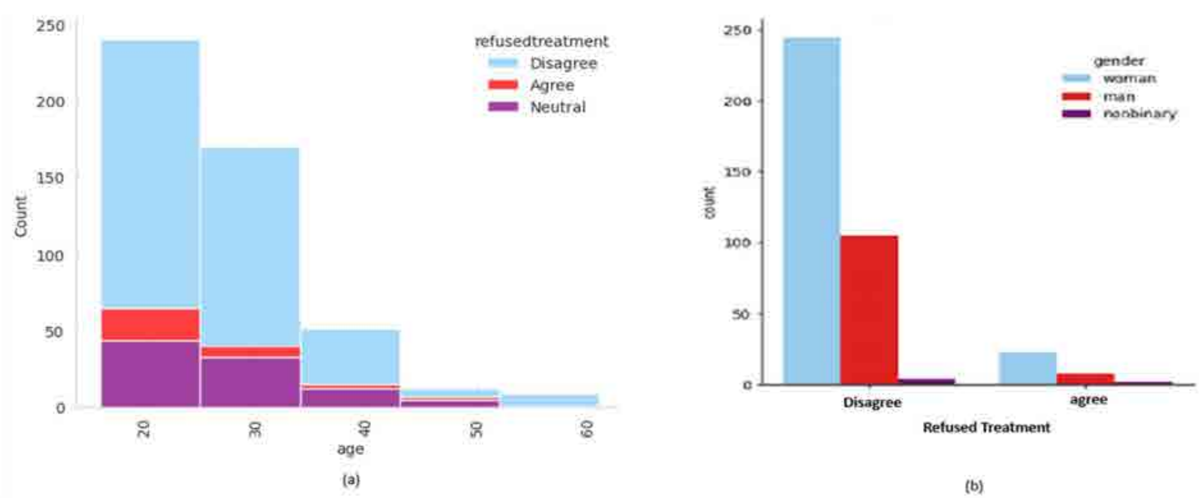

Figure 2. (a) Influence of Age (b) Influence of Gender on Healthcare Access.

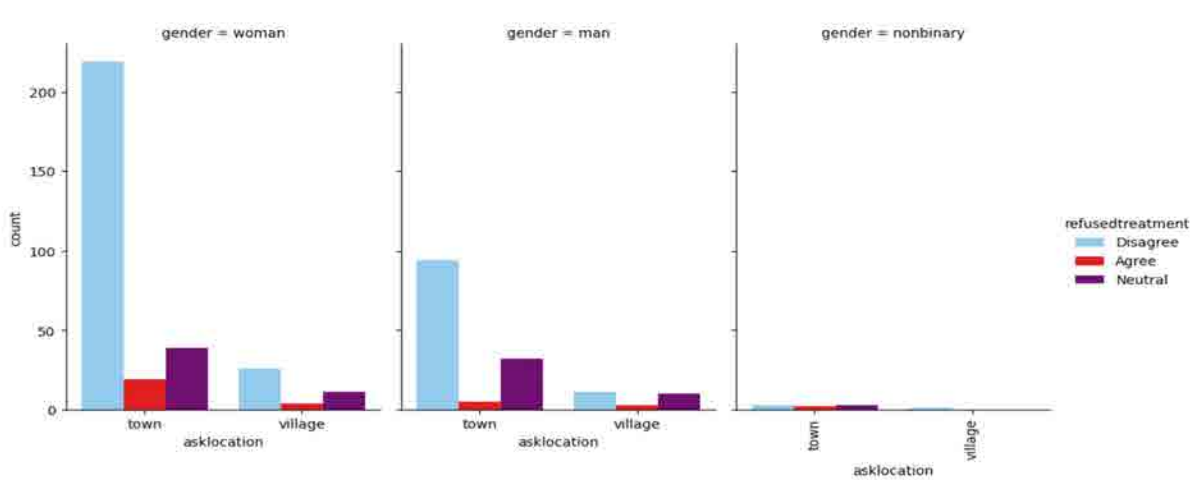

Figure 3. Impact of Gender on Healthcare.

healthcare with 10% of them (13 of 121) reporting being refused treatment. For the men, only 1 (of the 42 employed) and 3 (of the unemployed men) were refused treatment. However, like the women, a higher percentage (7%) of the self-employed men were refused treatment. There were 9 non-binary respondents and only 2 reported being refused treatment, with the first being unemployed and the other being employed. The drawn inference is that there seems to be a bias against self-employed persons, irrespective of gender. Statistically, a p-value of 0.656 for Employment and healthcare accessibility shown on Table 4, indicates that this is highly probable.

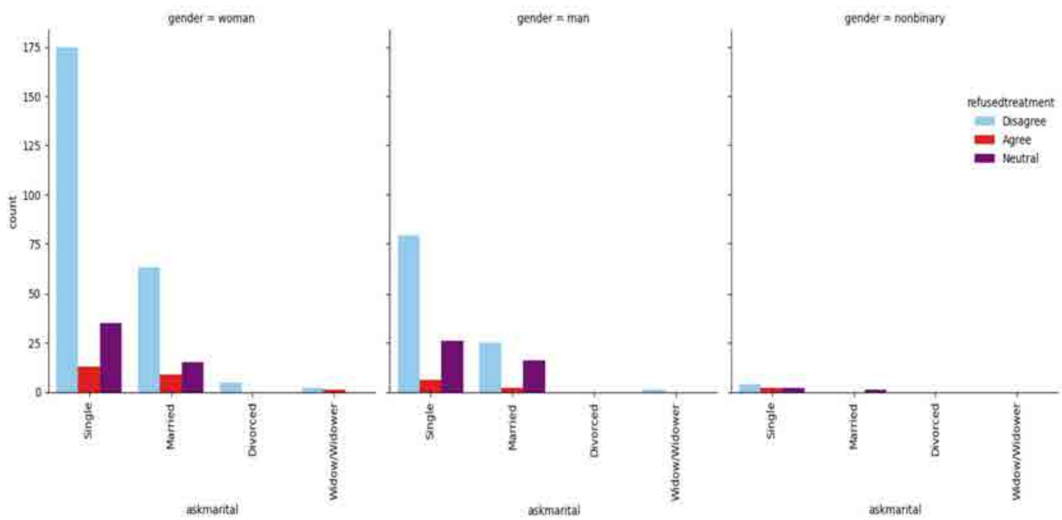

Figure 4. Influence of Marriage on Healthcare.

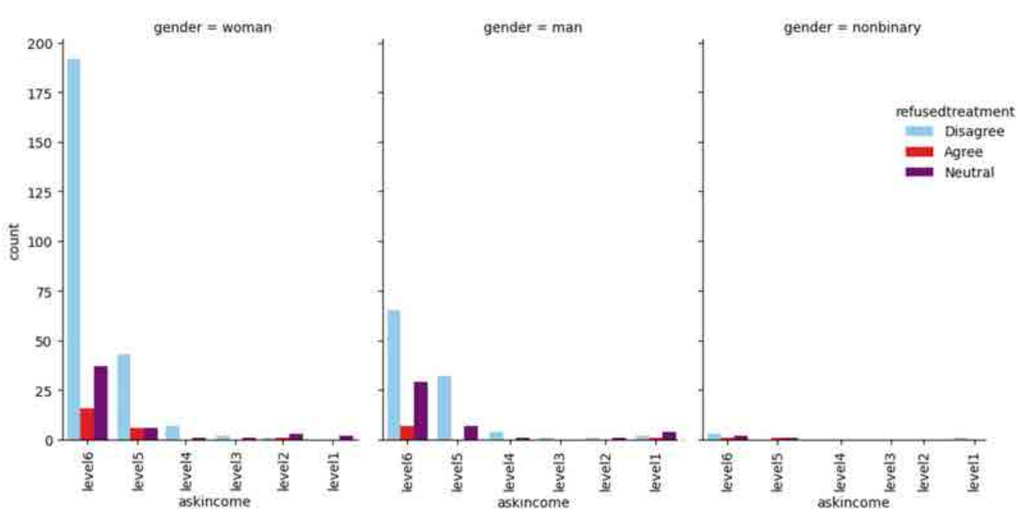

Figure 5. Influence of Income Levels on healthcare.

Table 5. Comparison of prediction results for raw & augmented data.

Model - class	Accuracy		Precision		Recall		F1-Score	
	Raw (%)	Augmt (%)	Raw (%)	Augmt (%)	Raw (%)	Augmt (%)	Raw (%)	Augmt (%)
CATBoost - 0	93	81	0	81	0	81	0	81
CATBoost - 1			92	81	100	81	96	81
SVC - 0	92	81	0	81	0	81	0	81
SVC - 1			92	81	100	81	96	81
RF - 0	93	81	100	77	19	87	32	82
RF - 1			93	85	100	74	97	79

Influence of education on gender-based healthcare accessibility

Figure 7 shows that among the 89 women with secondary level education and 194 with tertiary level education, 8 (8.9%) and 14 (7.2%) respectively reported being refused treatment. For the men, 6 (10%) of the 58 with secondary education and 1 of the 82 with tertiary level education were refused access to healthcare services. Similar trends were observed with the non-binary gender, as only those with secondary and/or tertiary level education were refused treatment. The drawn inference is that there seems to be a bias against higher levels of education irrespective of gender, but the bias seems more targeted at women with tertiary level education. Table 4 also buttresses this inference, with a weakly negative (-0.096) but statistically significant (p-value = 0.034) association between education and refused treatment.

Predictive data analysis

Using ML models, this section seeks to determine if an individual A, of G gender, marital status M, education level E, income level I, and shows symptoms S (*i.e.*, $P = A<G, M, E, I, S_x>$) would be able to access healthcare service.

Raw data

As earlier stated, the raw data was not balanced, with a ratio of 381:31 between the majority class (individuals who were able to access treatment) and minority class (those who were refused treatment). Table 5 shows a comparison of the prediction results for the minority class (0) using the raw and augmented datasets. On the table, "- 0" refers to the result obtained when the corresponding model was trained using the minority class data, while "– 1" refers to the result of the majority class. Using the raw data, RF yielded Recall and F1-score values of 19% and 32% respectively for the

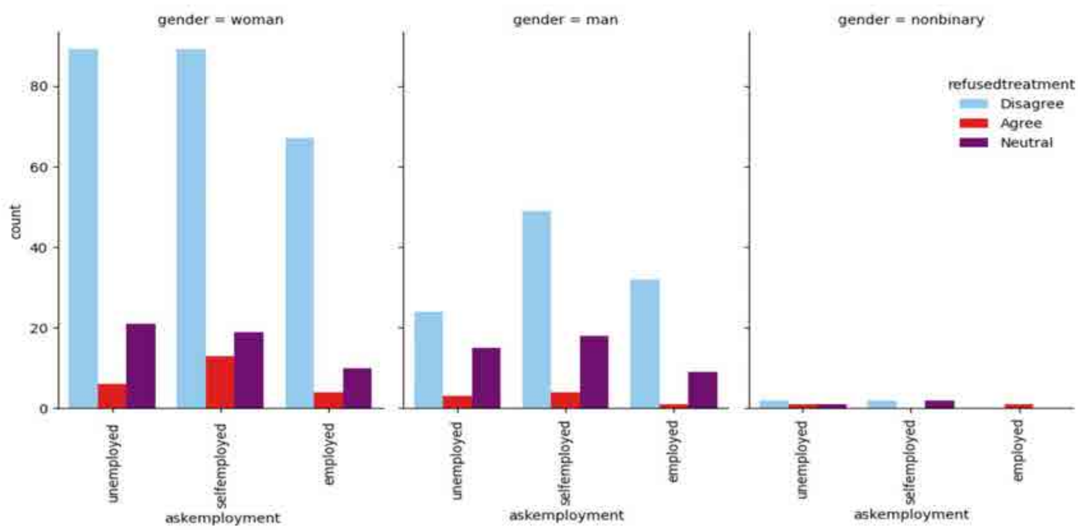

Figure 6. Influence of Employment on healthcare.

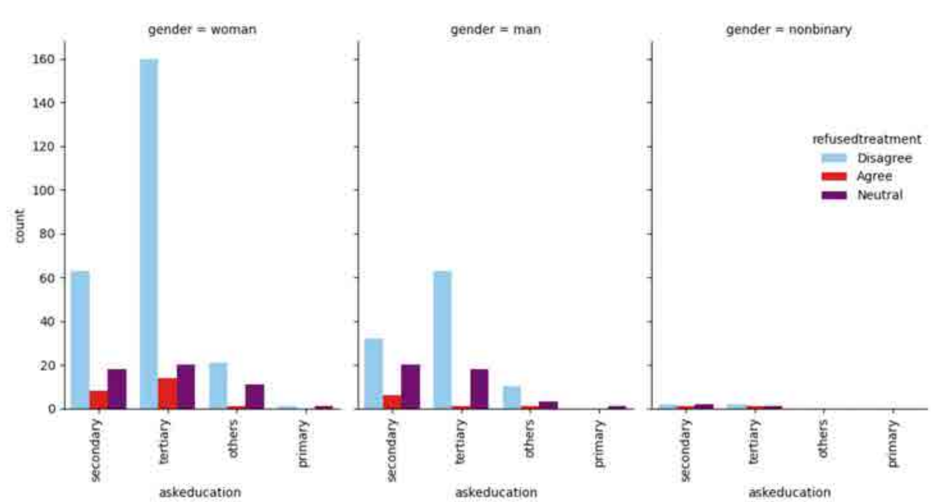

Figure 7. Influence of Education on healthcare.

minority class, while CATBoost and SVC both had 0% for Recall and F1-score. A value of 0 % implies that no classification was created for the minority class. This is in line with [16, 20] where Recall was reported as 0% for the minority class in highly imbalanced datasets. To address the poor prediction performance of the three models, we augmented the minority class data by oversampling it.

Augmented data

We augmented the minority class by oversampling it, then used the majority class (1) and augmented_minority_class data to train the three models. Obtained results, shown under "Augmt" columns on Table 5, reveal that the augmentation process significantly improved the prediction results of the minority class (0). The Recall and F1-Score values of CATBoost and SVC both rose to 81%, while RF's Recall rose from 19% to 87% and F1-Score rose from 32% to 82% for the minority class.

In summary, augmenting the data significantly improved the ML models' ability to predict healthcare accessibility. Therefore, the probability that an individual would be able to access (or be denied) healthcare services can be predicted with up to 81% accuracy on the average.

Discussions

From both analyses the following inferences can be drawn about factors that might influence accessibility to healthcare services: i) healthcare services are more accessible to the elderly; ii) women have a slight advantage in accessing healthcare services compared to the other genders; iii) women with lower income and education levels are marginally better favoured than those with higher income and education levels; the reverse is the case with men; iv) there seems to be a bias against self-employed persons irrespective of gender; v) with a balanced dataset, it is possible to predict, with up to 81% accuracy, if a patient with certain characteristics, socio-economic profile, and symptoms would receive treatment or not.

Assumptions and limitations

There were certain limitations and challenges encountered during this study, including: i) skewed dataset – there were more women respondents than other genders in the dataset; ii) the data was also highly imbalanced w.r.t. the 'refusedTreatment' attribute, as most of the respondents disagreed with being refused treatment, while only a few agreed; iii) the NER model used for symptom extraction did not always accurately identify symptoms in text, especially rare medical terms or misspelt words, e.g., it classified 'dizzyness' and 'dizziness' as different symptoms; iv) assumptions were made while grouping the symptoms together to get the top N; for instance, we grouped 'headache' and 'migraine' together because we assumed they are related; this might not always be correct, as they might be caused by or associated with completely different ailments.

Ethical considerations

Sex and gender

This work followed the Sex and Gender Equity in Research (SAGER) guidelines,[21] hence, sex refers to the biological and physiological characteristics of a person, and it is assigned at birth. Gender refers to the socially constructed characteristics of women and men, such as roles and norms. For this work, three gender identities were considered - men, women, and non-binary.

Marginalization

Individuals who fall into certain groups might be marginalized and excluded from accessing healthcare services. These include women, individuals living in rural areas, those with disabilities, illiterates, and single mothers. To mitigate this, we followed the P20 guidelines of the Development Initiative's Inclusive Data Charter (IDC),[22] which offers guidelines for identifying and analysing individuals that might be at the greatest risk of marginalisation.

Anonymity and privacy

Privacy concerns might dissuade people from giving accurate information. For example, women may give false information about their marital status, because marriage is generally considered a socially desirable status in Nigeria. To alleviate these concerns this research collected data anonymously.

Conclusions

Though considerable progress has been made in terms of healthcare accessibility globally, a lot still needs to be done to achieve universal health coverage, especially in rural areas of the world. In many remote and developing countries, accessible and affordable healthcare services remain elusive, while women are severely marginalised. Factors responsible for these include level of education, poverty, socio-cultural and religious beliefs.

This work considered regions in northern Nigeria and analysed data on healthcare accessibility, through individual, socio-economic, and equity lenses. Descriptive and predictive analyses were carried out on the data using statistical and Machine Learning (ML) models respectively. For the descriptive analysis, Pearson and Spearman's correlation coefficients were used, while CATBoost, Random Forest and Support Vector Machine ML models were used for the predictive analysis. The descriptive analysis revealed that women with lower income and education levels, and the elderly have a higher chance of accessing healthcare services. For predictive analysis, when an individual's biological characteristics, socio-economic status and medical symptoms are considered, ML models can predict the individual's accessibility to healthcare services, with up to 81% accuracy.

Data used in this work were in English language, which is not the native language in the region considered. Analysis of data in indigenous languages might be an avenue to extend this work. Additionally, exploration of other data augmentation techniques and the use of deep learning models for prediction might also be another direction for future work. Finally, other methods of quantifying qualitative data, such as fuzzy logic, could be considered for data pre-processing in the future.

References

1. United Nations. The sustainable development goals report 2023: Special edition. unstats.un.org. United Nations; 2023. Available from: https://unstats.un.org/sdgs/report/2023/
2. WHO. World bank and WHO: Half the world lacks access to essential health services, 100 million still pushed into extreme poverty because of health expenses. World Health Organization. World Health Organization; 2017. Available

from: https://www.who.int/news/item/13-12-2017-world-bank-and-who-half-the-world-lacks-access-to-essential-health-services-100-million-still-pushed-into-extreme-poverty-because-of-health-expenses
3. Top M, Konca M, Sapaz B. Technical efficiency of healthcare systems in African countries: an application based on data envelopment analysis. Health Policy Techn 2020;9:62-68.
4. Mupara L, Mogaka J, Brieger W, Tsoka-Gwegweni J. Scorecard metrics for assessing the extent of integration of community health worker programmes into national health systems. African J Primary Health Care Fam Med 2021;13:e1-e14.
5. Nicholas S. Health care equity and justice scorecard to increase diversity in clinical trial recruitment and retention. J Am Soc Nephrol 2022;33:1652-55.
6. Dhillon A, Singh A. Machine learning in healthcare data analysis: a survey. J Biology Today's World 2019;8:1-10.
7. Wongvibulsin S, Wu K, Zeger S. Clinical risk prediction with random forests for survival, longitudinal, and multivariate (RF-SLAM) data analysis. BMC Med Res Methodol 2020;20:1-14.
8. Cirillo D, Catuara-Solarz S, Morey C, et al. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. NPJ Digital Med 2020;3:1-11.
9. Straw I, Wu H. Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. BMJ Health & Care Informatics 2022;29:100457.
10. Chung H, Park C, Kang WS, Lee J. Gender bias in artificial intelligence: severity prediction at an early stage of COVID-19. Front Physiol 2021;12:2104.
11. Nivi. The Asknivi Product. Nivi. Available from: https://www.nivi.io/
12. Yinka-Banjo C, Akinyemi M, Ajayi O, Tresner-Kirsch D. Data on gender-equitable healthcare accessibility in northern Nigeria. Data Brief 2023;109979.
13. Neumann M, King D, Beltagy I, Ammar W. ScispaCy: Fast and Robust Models for Biomedical Natural Language Processing. Proceedings of the 18th BioNLP Workshop and Shared Task 2019;319–27.
14. Tresner-Kirsch D, Yinka-Banjo C, Akinyemi M, Ajayi O. Auditing natural language processing for gender equality in sub-Saharan African healthcare systems: framework development and evaluation. JMIR Preprints. doi:10.2196/preprints.51954.
15. Rovetta A. Raiders of the lost correlation: a guide on using Pearson and Spearman coefficients to detect hidden correlations in medical sciences. Cureus 2020;12:e11794.
16. Mohammed R, Rawashdeh J, Abdullah M. Machine learning with oversampling and undersampling techniques: overview study and experimental results. In2020: 11th International Conference on Information and Communication Systems (ICICS) 2020 Apr 7, IEEE. pp 243-48.
17. Prokhorenkova L, Gusev G, Vorobev A, et al. CatBoost: unbiased boosting with categorical features. Adv Neur In 2018;31:1-13.
18. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. J Mach Learn Res 2011;12:2825-30.
19. Ajayi O, Yinka-Banjo C, Akinyemi M, et al. Insights into gender-equity in healthcare accessibility in northern Nigeria [Source Code]. Code Ocean 2024. doi:10.24433/CO.2247195.v2
20. Gu Q, Zhu L, Cai Z. Evaluation measures of the classification performance of imbalanced data sets. In: Computational Intelligence and Intelligent Systems: 4th Int Symposium, ISICA 2009, Huangshi, China, 2009 Oct 23-25. Proceedings 4. Springer Berlin Heidelberg; 2009. pp. 461-471.
21. Heidari S, Babor T, De Castro P, et al. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 2016;1:2.
22. Development Initiatives. Systematic ways to 'leave no one behind'. n.d. [Internet]. Available from: https://www.data4sdgs.org/sites/default/files/file_uploads/JN_1286_IDC_KP_DI_CaseStudy.pdf