Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
 ISSN 1118 – 1931

PREDICTIVE HOSPITAL SITE SELECTION MODEL USING MACHINE LEARNING TECHNIQUES

Okengwu[1], U. A., Memmert[2], D., Rein[2], R. and Osuigbo,[3] E. N.
[1]Department of Computer Science, University of Port Harcourt, Port Harcourt, Nigeria
[2]Institute of Exercise Training and Sport Informatics, German Sport University, Cologne, Germany
[3]Department of Computer Science, Kenule Beeson Saro-Wiwa Polytechnic, Bori, Nigeria
[*]Corresponding Author: Email:ugochi.okengwu@uniport.edu.ng

Received: 09-02-2022
Accepted: 20-03-2022

ABSTRACT

Globally, countries are faced with healthcare challenges that vary from one to the next. While health service delivery challenges are more often seen in countries with a very high HumanDevelopment Index (HDI), inefficient healthcare intervention challenges attract more attention in those with a low HDI. Health systems and infrastructure interventions are major challenges for most countries in Africa. The conventional or traditional approach to situating hospitals has been subjected to the unreliable intuition of experts and perhaps biased by nepotism, favoritism, and tribalism of recognized interest. In this research, we prioritize health factors for hospital site predictions. A hospital is a healthcare intervention infrastructure that should meet the healthcare needs of people where it is located. Many hospital site selection researchers have considered other important factors such as geographical, economic, and socio-demographic factors. However, healthcare factors that will specifically address the healthcare needs of people in that locality have not been mentioned. This paper considers a robust, viable, reliable, and dependable approach to solve the specific problem of selecting the best location for building new hospitals based on the specific healthcare needs of the location for improved healthcare service delivery. We propose a supervised machine learning approach to predict the most suitable sites for building new hospitals. Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Logistic Regression (LGR), and Decision Tree (DT) machine learning algorithm are used to predict the optimal location for hospital site selection on the basis of various attributes used in the dataset. The dataset was extracted from clinical exploratory data using parameters such as reproductive, maternal, newborn, and child health, Infectious diseases, Non-communicable diseases, Adequate sanitation, Service capacity, and access to essential medicines as a Universal Health Coverage (UHC) standard by World Health Organization (WHO). The model is implemented using Python programming language. The system will improve health systems and infrastructural interventions, thereby enabling efficient healthcare service delivery in developing countries, especially in Africa.

Keywords: Decision Tree (DT), Healthcare, Hospital site selection, K-Nearest Neighbours (KNN) Logistic Regression (LGR) Prediction, Support Vector Machine (SVM).

INTRODUCTION

Man-made difficulties have harmed Africa's healthcare systems over the years, affecting institutional, human resource, financial, technical, and political advancements (Roncarolo et al., 2013). The majority of low-HDI countries, particularly in Africa, are unable to achieve the minimum criterion for good healthcare systems and infrastructure. In resource-constrained countries, inadequate service integration is connected to poor governance and human resource issues (Petersen et al., 2017; Marais and Petersen, 2015).

One key issue that has persisted over time is healthcare system inefficiency and infrastructure interventions, both of which have had a significant impact on the quality of healthcare service delivery. Cost, land availability, population, cultural concerns, manpower, proximity to man and natural elements, and access to inexpensive capital have all contributed to the difficulty of locating hospitals and healthcare institutions in various locations (Cynthia and Debra, 2011; Bob, 2012). These fundamental difficulties have had a significant impact on the availability of hospitals and medical centers for the general public. Because of the limited financial resources available to executive and administrative officers, further amenities have been completely neglected. These cost difficulties have also been exacerbated by a managerial or administrative officer's self-centered mindset, which has resulted in corrupt actions aimed at emptying public treasuries for personal gain. Although the land may appear to be available, due to the required size for establishing such hospitals, land availability can be a barrier. This can be tough to achieve in big cities with the required population size for siting hospitals, but less difficult in rural areas missing the fundamentals criteria fostering the siting of such establishment. The primary goal of hospital facilities is to provide health services within a specific geographic area in order to meet the healthcare needs of individuals (population size) who lack these services or who require more services. As a result, the population is critical in determining where hospitals should be located. The eventual location of medical institutions and their operational effectiveness might be determined by the cultural backdrop and background of a given community (Owen and Lawrence, 1996).

If hospital facilities are to function successfully, meeting the needs and desires of healthcare recipients, these elements cannot be overlooked. In fact, combining these criteria results in facilities that are strategically located to generate relevant money while providing essential services to the general public. As a result, the World Health Organization (WHO) developed a Universal Health Coverage (UHC) medical care standard to ensure that all people and communities have access to the health care they require without facing obstacles such as financial hardship, nepotism, or outright disregard.

In this study, an improved model that efficiently predicts optimal locations for situating hospitals using a supervised machine learning approach is developed. An efficient approach for situating the hospitals is undoubtedly a predominant challenge that has affected the quality of health service delivery in developing countries. Health systems and infrastructure are crucial components of a wider healthcare intervention approach, but if not placed correctly, they can disrupt the entire process. The conventional or traditional approach to situating hospitals has been subjected to the unreliable intuition of experts and perhaps biased by nepotism, favoritism, and tribalism of recognized interest. In this work, we propose a robust and efficient model for situating hospitals in optimal locations for

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

improved healthcare service delivery. This approach will apply supervised machine learning algorithms using Support Vector Machine (SVM), K-Nearest Neighbours (KKN), Logistic Regression (LGR), and Decision Tree (DT) and select the best model for implementation.

LITERATURE REVIEW

Site Selection

Site selection is a research problem whose purpose is to find the best location that meets a set of criteria (Senvar et al., 2016). As proposed by Mohammed et al., site selection typically includes two primary stages: screening (identifying a set number of prospective sites from a wide topographical region given a scope of selection elements) and assessment (extensive evaluation of options to determine the most appropriate site) (2019). The two steps involve a large number of sometimes problematic factors. When you're going through the same situation, there are a variety of tools available to help you choose the best location. Expert Systems (ES) are used for well-defined and organized issues, whereas Decision Support Systems (DSS) are used for poorly organized situations or a combination of the two((Vafaei and ztayşi, 2014).

GIS with MCDM methodologies that can help with site selection in instances when the problem is poorly organized, meaning that leaders lack complete and reliable data on specifications, options, and outcomes (Boyac and Işman, 2021). Traditional GIS site selection methods rely on the transformation of powerful layers into a classified map, such as through the use of a Boolean model or Index Overlay activity.

Various hospital site selection scholars have embraced the comprehensive assessment of a hospital site and proposed a wide range of site selection standards that are largely dominated by topographical, financial, and socio-segment variables, with little regard for medical services needs evaluation. Separation from the arterial road, travel duration, tainting, land cost, and demographics were all factors evaluated by Vahinia et al. (2008). Soltani and Marandi (2011) looked at the distance to major streets, distance to other clinical facilities, demography, and land size. Wu et al. (2007) analyzed population number, age, thickness, legislative methods, capital, work, and land while deciding on the best location for Taiwanese hospitals. Schuurman et al. (2006) also looked at the importance of the help region's socioeconomics and proximity to potential expansion, space-separated from movement time, and populace thickness.

In light of the lack of available literature addressing healthcare characteristics as a significant measure for selecting hospital sites, this study considers four significant qualities and additional sub-traits for analyzing the medical care requirements of a competitive clinic site in Rivers state. These factors are based on the World Health Organization's Universal Health Coverage (UHC) - Service Coverage Indicators (SCI) standard (WHO). UHC, according to Bloom et al. (2018), means that everybody can get the healthcare they need, when and where they need it, regardless of their financial situation.

Healthcare Interventions

The word "health intervention" refers to any activity taken with the goal of improving human health by preventing disease, reducing the duration of an existing infection, or restoring capacity lost due to illness or injury (Glasziou et al., 2010). A wide range of new interventions, as well as new approaches for using mediations, are being developed to combat the major diseases that plague low- and middle-income countries. Some of these treatments cover both public health and clinical considerations, and include medications for acute and chronic conditions, antibodies, vector control, health education, behavior change systems, injury prevention, and better wellbeing planning and the board strategies that work on a variety of wellbeing-related activities (Clarke et al., 2019). Interventions might be best applied in populaces to decide their effect on working on the health of the populace utilizing a field trial assessment (Smit et al., 2020. Interventions can be grouped into two general classes:

i. Preventive interventions are those that keep the sickness from happening and in this manner diminish the frequency (new instances) of infection, vaccines, nutritional interventions, maternal and neonatal interventions, education and conduct change, environmental adjustments, vector, and transitional host control, medication for the avoidance of sickness, and Injury anticipation are on the whole instances of preventive medical care mediations (Clarke et al., 2019).
ii. Therapeutic interventions are those that treat, moderate, or delay the impacts of sickness when it is in progress, and in this way diminish the case casualty rate or lessen the inability or dismalnessconnected with a disease. Health systems and infrastructure interventions are all examples of therapeutic interventions (Clarke et al., 2019).
iii. A few interventions might make the two impacts. Other types of interventions are legislation, legitimate activity, tax collection, and sponsorships, Implementation research, and Complex interventions.

Universal Health Coverage (UHC) Framework

Universal Health Coverage (UHC) is a framework designed by the World Health Organization (WHO) to ensure that all individuals and networks have access to the healthcare they require without facing financial hardship. It encompasses the complete range of critical, high-quality health services, from prevention through treatment, rehabilitation, and palliative care for people of all ages (Bloom et al., 2018). These administrations want enough and capable healthcare practitioners with optimal abilities blend at the office, outreach, and community level, who are impartially disseminated, satisfactorily supported, and appreciate good work. UHC systems enable everyone to access health services that address the primary causes of illness and death, while also ensuring that the quality of those services is sufficient to improve the health of people who receive them. Protecting people from the financial consequences of paying for health services out of their own pockets reduces the risk that they will be forced into poverty because unexpected disease forces them to burn through their life investment funds, sell assets, or acquire - obliterating their prospects and, in many cases, those of their loved ones. Accomplishing UHC is one of the objectives the countries of the world set

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

while taking on the SDGs in 2015 (Xu et al., 2015).

At the United Nations General Assembly High-Level Meeting on UHC in 2019, nations reiterated their commitment to achieving UHC. Nations agreed that UHC will make progress toward other well-being goals. Good health enables children to learn and adults to purchase support people in escaping poverty and provides the foundation for long-term financial success (Nygren-Krug, 2019).

WHO contributes to achieving the Thirteenth General Program of Work 2025 goal of 1 billion more people benefiting from UHC, while also contributing to the two other billion goals of 1 billion more people better protected from health crises and 1 billion more people benefiting from better health and prosperity? It also contributes to WHO's fundamental goal of the right to the most noteworthy achievable norm of wellbeing, to Health for All, and the SDGs.

Parameters for Measuring UHC

Checking the advancement towards UHC is conceivable and ought to zero in on two things:

i. The extent of a populace that can get to fundamental quality healthcare services (SDG 3.8.1).

ii. The extent of the populace that spends a lot of family income on healthcare (SDG 3.8.2).

Estimating value is additionally basic to comprehend who is as a rule left behind-where and why. Along with the World Bank, WHO has fostered a structure to follow the advancement of UHC by checking the two classifications, considering both the general level and the degree to which UHC is evenhanded, offering administration inclusion and monetary security to all individuals inside a populace, like poor people or those living in distant provincial regions. WHO involves sixteen fundamental healthcare services in four classes as signs of the level and value of inclusion in nations as displayed in Table 1:

Some data, like the number of pages of an article or catchphrases, some of the time didn't show up in the StArt tool. Hence, the analysts needed more data to decide if the article would be incorporated or avoided in stage 2. Subsequently, analysts investigated these articles physically in stage 3. In this stage, a few articles, for example, short papers and articles not written in English were found.

Table 1: Universal Health Coverage Attributes.

Attributes	**Sub-attributes**
Reproductive, maternal, newborn and child health (RMNC)	Family Planning Antenatal care Full child immunization Health-seeking behaviour for child illness
Infectious diseases (ID)	Malaria prevention and treatment Tuberculosis effective treatment

	Cholera Prevention and treatment HIV antiretroviral treatment Adequate sanitation
Noncommunicable diseases (ND)	Prevalence of raised blood pressure Mean fasting plasma glucose Cervical cancer screening Tobacco control
Service capacity and access (SCA)	Basic hospital access Health worker density Access to essential medicines Compliance with the International Health Regulations

CATEGORY 1: Reproductive, maternal, newborn and child health:

i. family planning
ii. antenatal and delivery care
iii. full child immunization
iv. health-seeking behaviour for pneumonia.

CATEGORY 2: Infectious diseases:

i. tuberculosis treatment
ii. HIV antiretroviral treatment
iii. use of insecticide-treated bed nets for malaria prevention
iv. adequate sanitation.

CATEGORY 3: Noncommunicable diseases:

i. prevention and treatment of raised blood pressure
ii. prevention and treatment of raised blood glucose
iii. cervical cancer screening
iv. tobacco (non-) smoking.

CATEGORY 4: Service capacity and access:

i. basic hospital access
ii. health worker density
iii. access to essential medicines
iv. health security: compliance with the International Health Regulations.

Every nation is interesting, and every nation might zero in on various regions, or foster their own specific manners of estimating progress towards UHC. In any case, there is likewise esteem in a worldwide methodology that utilizations normalized measures that are globally perceived so they are practically identical across borders and over the long run..

Study Area Evaluation

The area under investigation covers all Local Government Areas (LGAs) in Rivers State, an oil rich state located in the southern part of Nigeria. The area extends from 51023’ to 51029’ east and from 35044’ to 35048’ south. The study area and existing hospitals are shown in Fig 1. The research evaluates the healthcare needs of the study area and predicts optimal locations for new hospital sites based on the health requirements of the selected location.

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

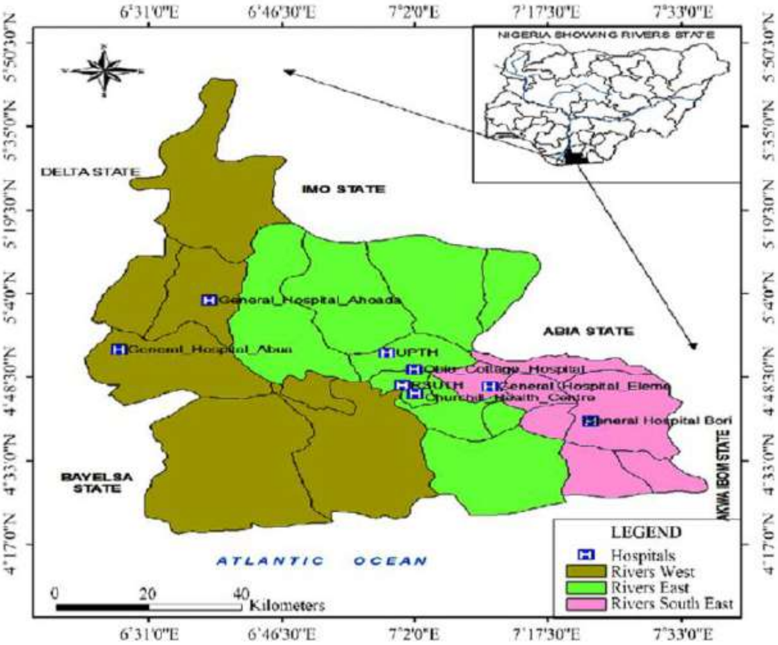

Fig. 1. Map of Rivers State showing existing hospitals.

METHODOLOGY AND ANALYSIS

The Rivers State Hospital dataset has fields without values and some values are more common than other values in the field, which makes the data distribution of the dataset unbalanced. In order to enhance the predictive performance of the model, a predictive machine learning method is proposed in this paper. The method consists of two phases: Data Preprocessing and Prediction. The whole process used in the methodology is shown in Fig 2.

Data Preprocessing

In the data preprocessing phase, data is cleaned and encoded for efficient performance of the predictive model. The data preprocessing approach used in this paper consist of three steps: Data cleaning, Value weighting and encoding.

i. *Data Cleaning:* This first step is performed to identify and correct errors in the dataset that may affect the prediction model. Columns that have a single value are probably useless for modelling and are considered as zero-variance predictors. Depending on the choice of modeling algorithm used, variables with a single value can also cause errors or unexpected results. Some columns referred to as near-zero variance predictors have few unique values (very small number close to zero) and occur infrequently in the data. Although near-zero variance predictors likely contain little valuable predictive information, they were still considered for prediction. To highlight these columns, we calculate the number of unique values for each variable as a percentage of the total number of rows in the dataset.

Visualization of Cleaned Data: The purpose of data cleaning process is to remove columns that don't contain much information and duplicated rows to reduce the imbalance of dataset. Fig 3, 4, 5 and 6 shows the data visualization for the different parameters in the collected data indicating how they are distributed in some Local Governments Areas in Rivers State. Table 2 shows the Local Government Areas and the important variables for prediction.

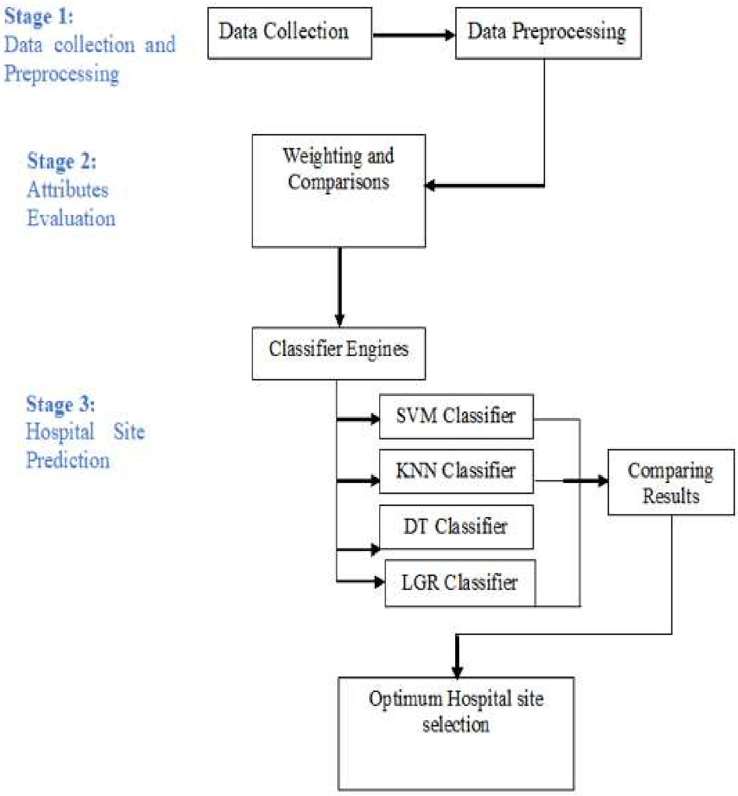

Fig. 2. System flowchart of the methodology used in this study.

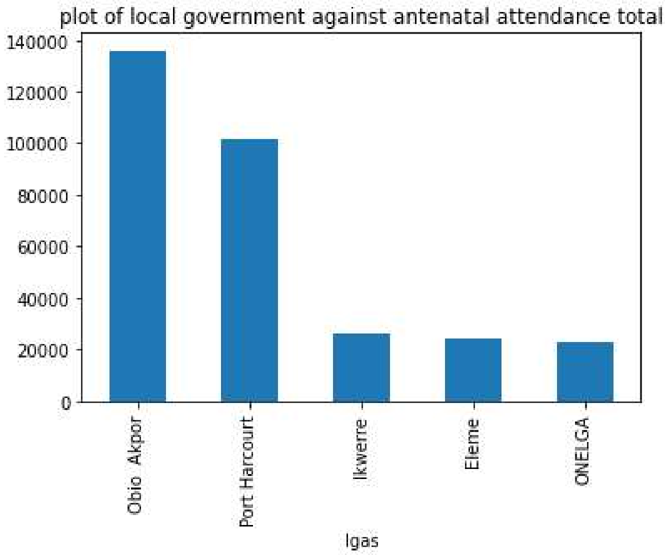

Fig. 3. Antenatal attendance across the Local Governments

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

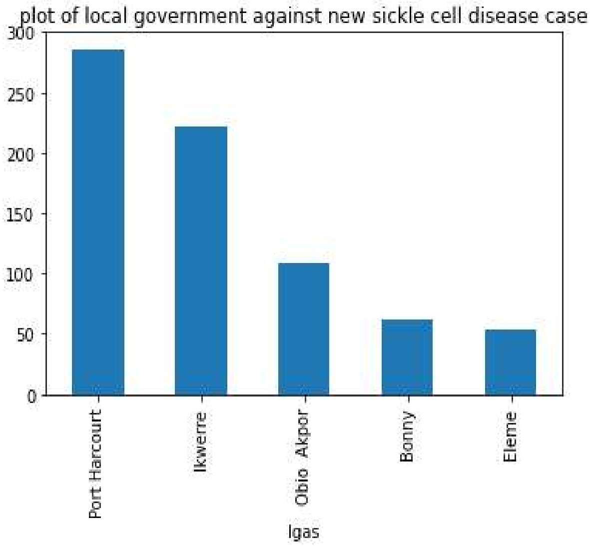

Fig. 4. HIV positive patients identified across the Local Governments

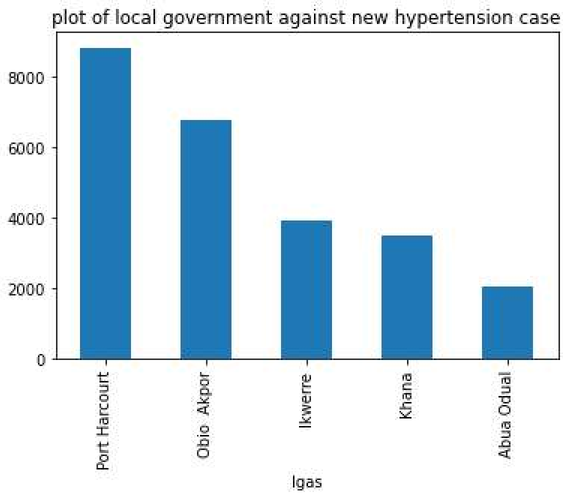

Fig. 5. Hypertension cases identified across the Local Governments

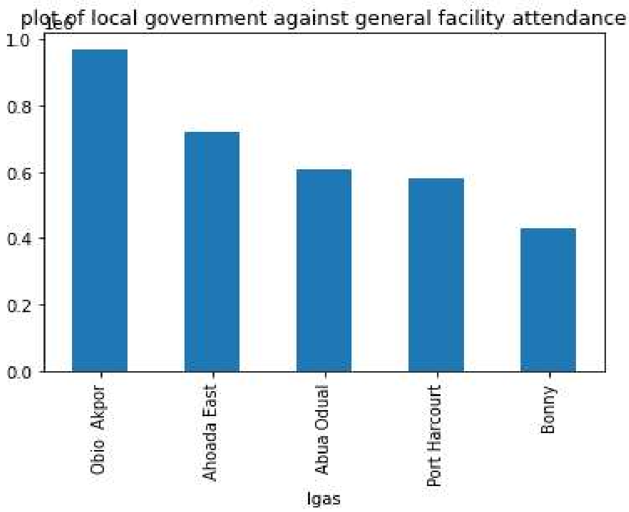

Fig. 6. The rate of attendance of patients to the hospitals in the Local Governments

Table 2: Important Variables for Prediction.

	Antenatal Care	HIV Positive	Hypertension	Hospital Facility Access
ObioAkpor	136061	58470	6775	971100
Port Harcourt	101533	33861	8832	579920
Khana	170	29166	3500	94309
Ikwerre	26019	364	3891	431265
Eleme	24193	28252	554	85407

Table 3: Feature importance score for the highest variables

UHC Category	Feature	Importance
Infectious diseases (ID)	HIV Positive	0.079
Reproductive, maternal, newborn and child health (RMNC)	Antenatal Care	0.056
Noncommunicable diseases (ND)	Hypertension	0.037
Service capacity and access (SCA)	Hospital Facility Access	0.037

ii. *Value Weighting:* Next, value weighing is performed as part of the data preprocessing phase. Value weighting deals with the fact that particular values in a field might be more common than other values in that field. The Machine Learning Algorithm used in this paper is supervised learning algorithm, so predicted output was derived based on the parameters or variables available by calculating percentage mean:

= [Sum of %n variables across each local government / Total number of variables]

The coincidence of rare values in a field adds more to the overall similarity than the coincidence of frequent values. A set average weight was given to be 0.75 based on the range of the values. A set average weight to each value according to its probability in the input data was given to be 0.75 corresponding to the range of the values. Rare values carry a large weight, while common values carry a small weight. This weight is used for both matching and non-matching records and determines where hospital can be located. Table 3 shows the feature importance score for highest values.

iii. *Encoding:* In this step, LGA variable is converted to numerical values by using label encoding in python as shown in Fig. 7, to enable the machine learning algorithms used in the predictive model to handle these variables. Most machine learning algorithms cannot handle non-numerical variables.

iv. *Feature Selection:* Most useful features which will be used to train the prediction model are selected in this step. Dimensionality reduction is used to remove features features that are not contributing enough to the overall variance. Important feature function was used to reduce the features to a shape of 5 rows and 124 columns as shown in Fig.7. There are different ranges of values for all variables and feature

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ISSN 1118 – 1931

scaling was used to convert those values and have them under the same range of values.

The purpose of feature selection is to find the best feature subset that can maximize the performance of the prediction. Once the optimal training dataset and the optimal feature subset are selected, those will be taken into the Prediction phase.

Prediction Phase

In this phase, the goal is to predict the optimal site to build a new hospital with a suggestion of the type of hospital. Four machine learning algorithms Support Vector Machine (SVM), K-Nearest Neigbours (KNN), Decision Tree (DT) and Linear Logistic Regression (LGR). Fig. 8 shows the prediction hierarchical structure for the model decision-making with S1, S2, S3, and S4 selected as candidate sites for prediction. The attributes considered for the decision making are seen in Table 3. The classifiers were trained and tested using the dataset with the split ratio 75% for training data and 25% for the test data. Accuracy metrics were measured and confusion matrix of each model was calculated to validate the best model for implementation as shown in Fig. 9, 10, 11 and 12.

	general facility attendance	general outpatient attendance	antenatal attendance total	deliveries-assisted	caesarean section	deliveries complications	deliveries normal	preterm deliveries	pregnacy outcome-birth asphyxia	pregnacy outcome--neonatal-jaundice	...	individuals started on tb treatment ((hiv-ve)
0	114760.0	24053.0	4773.0	6.0	19.0	2.0	63.0	1.0	3.0	6.0	...	188.0
1	21936.0	4869.0	1626.0	6.0	19.0	2.0	63.0	1.0	3.0	6.0	...	23.0
2	21936.0	4869.0	1626.0	6.0	19.0	2.0	63.0	1.0	3.0	6.0	...	23.0
3	21936.0	4869.0	1626.0	6.0	19.0	2.0	63.0	1.0	3.0	6.0	...	23.0
4	21936.0	4869.0	1626.0	6.0	19.0	2.0	63.0	1.0	3.0	6.0	...	2.0

5 rows × 124 columns

Fig. 7. Label encoder and transformation

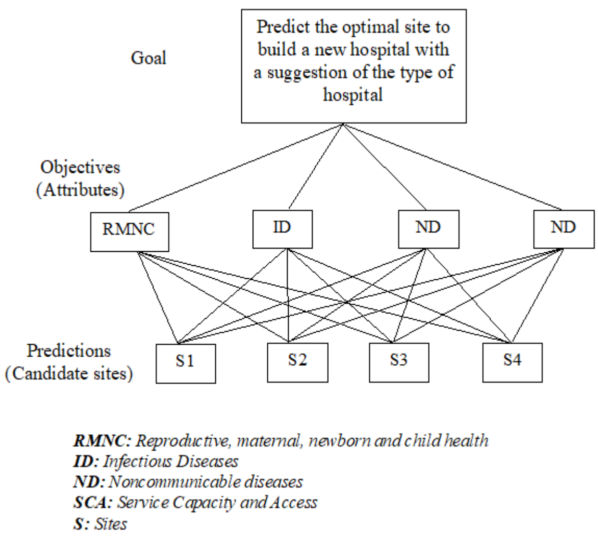

Fig. 8. Prediction hierarchical structure for decision-making

RESULTS

To test the model, an experiment was performed on a PC with Intel(R) Core (TM) i5-4460 at 3.6 GHz CPU and 8GB memory, running on Windows 10. Programs are coded in Python using PyCharm Professional Edition 2021.3.1 environment on the version of Anaconda 2020.11. Empirical values are used to obtain the parameters for the algorithm. The dataset used for the experiment was gotten from Rivers State Ministry of Health. Data was collected across all local government in Rivers State, Nigeria having a shape of 207 rows and 132 columns. Pairs of attributes are subjectively evaluated on a nine-point scale. The important healthcare attributes and sub-attributes considered for selecting optimal hospital sites as adopted from Universal Health Coverage are summarized in Table 1, these attributes and sub-attributes address the healthcare needs of different locations for siting hospitals. The dataset and codes for this study can be accessed on GitHub.

Four machine learning algorithms were used in the predictive model as seen in Fig. 2. To compare the performance of these algorithms and select the most efficient one for optimal hospital site selection, Fig. 9 shows the confusion matrix of the four algorithms used for prediction over the dataset. The results shows that Support Vector Machine (SVM) performed better compared to other algorithms for predicting the target variable. To verify the effectiveness of SVM for optimal hospital site selection model proposed in this paper, the precision, recall, and F1_score obtained as shown in Table 4.

Table 4: SVM performance on Precision, Recall, and F1-score

	Precision	Recall	F1-score	Support
0	0.98	1.00	0.99	47
1	1.00	0.80	0.89	5
Accuracy			0.98	52
Macro Avg	0.99	0.90	0.94	52
Weighted Avg	0.98	0.98	0.98	52

The results show the local governments where hospitals will not be built produced 0.98 and the local governments where hospitals will be built produced 1.00 in precision. Recall 1.00 for local governments where hospitals will not be built and 0.80 for local governments where hospitals should be built. For f1-score 0.99 for local governments where hospitals will not be built and 0.89 for local governments where hospitals will be built as classified by the proposed system model.

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 *https://dx.doi.org/10.4314/sa.v21i1.14*
© Faculty of Science, University of Port Harcourt, Printed in Nigeria ***ISSN 1118 – 1931***

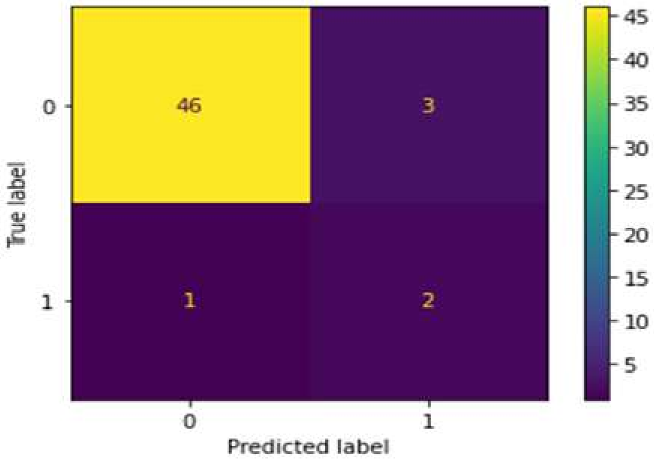

Fig. 9. Confusion matrix to show accuracy of SVM

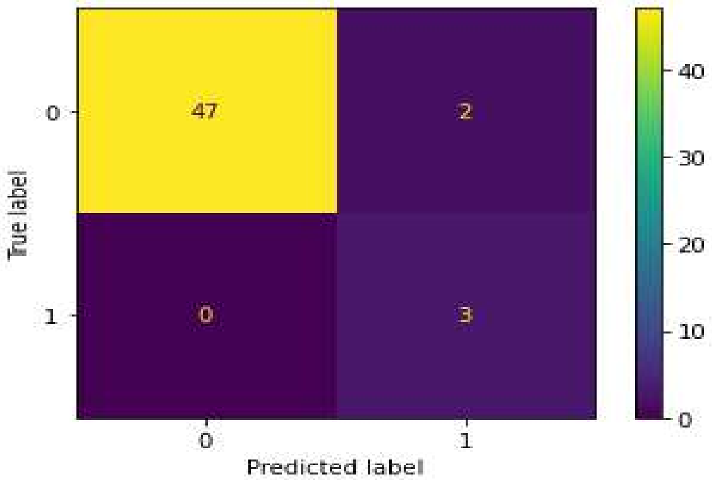

Fig. 10. Confusion matrix to show accuracy of KNN.

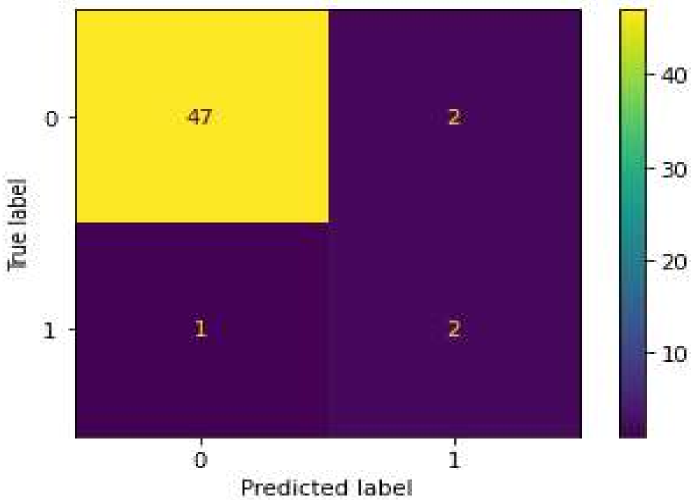

Fig. 11. Confusion matrix to show accuracy of LGR

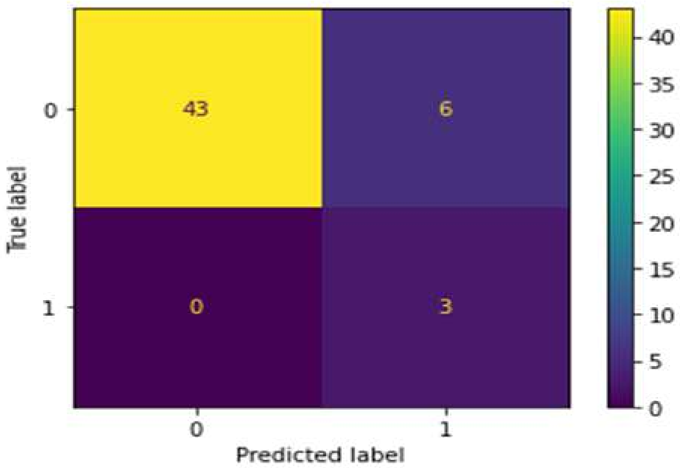

Fig. 12. Confusion matrix to show accuracy of DT

SVM showed accuracy rate of 96.15%, significantly higher compared to 94.23%, 92.31%, and 88.46% obtained by LGR, KNN, and DT classifiers respectively. It can be seen that SVM has achieved good performance on the prediction of optimal hospital site.

The Optimized Hospital Site Selection Predictive model is implemented with Support Vector Machine, due to high accuracy rate. The model is implemented with data trial and predicted that hospital can be built in Obi Akpor Local Government, while the model predicts that hospital should not be built in Emohua Local Government as shown in Figure 10.

CONCLUSION

Hospital site selection is a research problem common in developing countries with the goal to find the optimum location that satisfies some predefined criteria and give the populace of a particular locality have access to adequate healthcare services such as hospitals. Machine learning approach was applied using Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Logistics Regression (LGR), and Decision Tree (DT) as classifier engines to predict hospital sites.

In this study, the model determined the optimum site for hospital site in Rivers State, Nigeria. The 23 local government areas in the state are used as possible location parameters. First, Percentage Mean of the features was determined for the screening and selection of appropriate options based on the World Health Organization (WHO) Universal Health Coverage (UHC) standard parameters to generate the possible local government to site a new hospital. Infectious Disease (ID) features were the predominant features selected as important features. In the second phase, important features are classified using different classifier engines. SVM performed better with accuracy rate of 96.15% compared to the results of KNN with accuracy rate of 92.31%, LGR with accuracy rate of 94.23%. Decision Tree with accuracy rate of 88.46%.

The proposed prediction method can be used as a reference for medical service administrators to select the optimal location for a new hospital to ensure that it

Scientia Africana, Vol. 21 (No. 1), April, 2022. Pp 159-174 https://dx.doi.org/10.4314/sa.v21i1.14
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

meets the health care needs of people in a particular location using the Universal Health Coverage Standard. This can be considered as a valuable prototype and reference for hospital administrators and academics in establishing a standardized means of selecting location for medical care facilities. This study contributes healthcare intervention by suggesting a more robust decision-making tool for hospital site selection.

As a recommendation for further study on this research, more healthcare factors can be considered as criteria for alternatives. Some of potentially significant healthcare factors as suggested by World Health Organization (WHO) Universal Health Coverage (UHC) standard parameters and others like geographical, economical and socio-demographic factors affecting a hospital site selection were absent in the analysis. Also, a hybridized approach can be used combining the computing efficiencies of two or more machine learning techniques.

Acknowledgment

This study was sponsored by The World Academy of Science (TWAS) and German Research Foundation (DFG).

REFERENCES

Bloom, D. E., A. Khoury, and R. Subbaraman (2018). The promise and peril of universal health care. *Science*, 361(6404).

Bob H. (2012). 5 Challenges in Renovating or Building a Hospital, *retrieved online from http://www.beckershospitalreview.com/strategic-planning/5-challenges-in-renovating-or-building-a-hospital.html*, July 2017.

Boyacı, A. Ç., and A. Şişman (2021). Pandemic hospital site selection: a GIS-based MCDM approach employing Pythagorean fuzzy sets. *Environmental Science and Pollution Research*, 1-13.

Clarke, G. M., S. Conti, A. T. Wolters, and A. Steventon (2019). Evaluating the impact of healthcare interventions using routine data. *British Medical Journal*, 365.

Cynthia A. L., and Debra H. (2011). Design Details for Health: Making the Most of Design's Healing Potential. *2nd Edition John Wiley & Sons*, Inc.

Glasziou, P., I. Chalmers, D. G. Altman, H. Bastian, I. Boutron, and J. W. Williams (2010). Taking healthcare interventions from trial to practice. *British Medical Journal*, 341.

Marais D. L., I. Petersen (2015). Health system governance to support integrated mental health care in South Africa: challenges and opportunities. *International Journal Mental Health Systems*, 9(1), 14.

Mohammed, H., Z. Majid, Y. Yamusa, M. Ariff, K. Idris, and N. Darwin (2019). Sanitary landfill siting using GIS and AHP. *Engineering, Technology & Applied Science Research*, 9(3), 4100-4104.

Nsaif, Q. A., S. M. Khaleel, and A. H. Khateeb (2020). Integration of GIS And Remote Sensing Technique for Hospital Site Selection in Baquba District. *Journal of Engineering Science and Technology*, 15(3), 1492-1505.

Nygren-Krug, H. (2019). The right (s) road to universal health coverage. *Health and human rights*, 21(2), 215.

Okengwu, U. A., and E. N. Osuigbo (2021). Optimal Hospital Site Selection Software (Version 1.0) [Computer software]. https://github.com/ugodepaker/Riversstate_Hospital_Model

Owen B. H., and P. Lawrence (1996). The Planning and Design Process, 2nd Edition by Rockvillie Md,

Petersen I., D. Marais, J. Abdulmalik (2017). Strengthening mental health system governance in six low-and middle-income countries in Africa and South Asia: challenges, needs and potential strategies. *Health Policy Plan*, 32(5), 699–709.

Roncarolo, F., A. Boivin, J.L. Denis, R. Hébert, P. Lehoux (2013). What do we know about the needs and challenges of health systems? A scoping review of the international literature. *BMC Health Serv Res*, 17(1), 636.

Schuurman, N., R. S. Fiedler, S. Grzybowski, and D. Grund (2006). Defining rational hospital catchments for non-urban areas based on travel-time. *International Journal of Health Geographics*, http://www.ijhealth geographics.com/content/5/1/43 (Accessed on 7th September 2021).

Senvar, O., I. Otay, and E. Bolturk (2016). Hospital site selection via hesitant fuzzy TOPSIS. *IFAC-Papers On Line*, *49*(12), 1140-1145.

Smit, L. C., J. Dikken, M. J. Schuurmans, N. J. de Wit, and N. Bleijenberg (2020). Value of social network analysis for developing and evaluating complex healthcare interventions: a scoping review. *British Medical Journal*, 10(11).

Soltani, A., E. Z. Marandi (2011). Hospital site selection using two stage fuzzy multi-criteria decision-making process. *Journal of Urban andenvironmental engineering*, 5(1), 32-43.

Vahidnia, M.H., A. Alesheikh, and A. Alimohammadi (2009). Hospital site selection using fuzzy AHP and its derivatives. *Journal of environmental management*, 90(10), 3048-3056.

Wu, C.R., C. Lin, and H. Chen (2007). Optimal selection of Taiwanese hospitals to ensure a competitive advantage by using analytical hierarchy process and sensitivity analysis. Building and environment, 42(3), 1431-1444.

Xu, Y., C. Huang, and U. Colón-Ramos (2015). Moving Toward Universal Health Coverage (UHC) to Achieve Inclusive and Sustainable Health Development: Three Essential Strategies Drawn from Asian Experience: Comment on" Improving the World's Health Through the Post-2015 Development Agenda: Perspectives from Rwanda". *International journal of health policy and management*, 4(12), 869.

Yang, J. and P. Shi (2002). Applying Analytic Hierarchy Process in Firm's Overall Performance Evaluation: A Case Study in China. *International Journal of Business.* 7(1), 30-46.