Journal of Data Analysis and Information Processing, 2022, 10, 170-183
https://www.scirp.org/journal/jdaip
ISSN Online: 2327-7203
ISSN Print: 2327-7211

Maternal and Child Health Care Quality Assessment: An Improved Approach Using *K*-Means Clustering

Sarah Nyanjara[1], Dina Machuve[1], Pirkko Nykanen[2]

[1]Department of Communication Science and Engineering, Nelson Mandela Institution of Science and Technology, Arusha, Tanzania
[2]School for Information Sciences, Center for Information and Systems, University of Tampere, Tampere, Finland
Email: nyanjaras@nm-aist.ac.tz

How to cite this paper: Nyanjara, S., Machuve, D. and Nykanen, P. (2022) Maternal and Child Health Care Quality Assessment: An Improved Approach Using *K*-Means Clustering. *Journal of Data Analysis and Information Processing*, **10**, 170-183.
https://doi.org/10.4236/jdaip.2022.103011

Received: June 27, 2022
Accepted: August 5, 2022
Published: August 8, 2022

Abstract

High maternal and child deaths in developing countries are frequently linked to poor health services provided to pregnant women and children. To improve the quality of maternal, neonatal and child health (MNCH) services, the government and other stakeholders in MNCH emphasize the importance of quality assessment. However, effective quality assessment approaches are mostly lacking in most developing countries, particularly in Tanzania. This study, therefore, aimed at developing a quality assessment approach that can effectively assess and report on the quality of MNCH services. Due to the need for a good quality assessment approach that suits a resource-constrained environment, machine learning-based approach was proposed and developed. *K*-means algorithm was used to develop a clustering model that groups MNCH data and performs cluster summarization to discover the knowledge portrayed in each group on the quality of MNCH services. Results confirmed the clustering model's ability to assign the data points into appropriate clusters; cluster analysis with the collaboration of MNCH experts successfully discovered insights on the quality of services portrayed by each group.

Keywords

Maternal Health Quality, Clustering Model, Health Quality Assessment, Maternal Health Assessment

1. Introduction

Focus on healthcare quality assessment had strengthened since the 1980s when Donabedian developed the first framework for quality measurement [1]. That framework focused on quality elements related to healthcare structure, process and outcome [2] [3]. Since then, various scholars have attempted to focus on and define quality in healthcare in different contexts [4]. Godlee [5] has defined *quality* as "clinically effective, safe and a good experience for the patient". The Institute of Medicine (IOM) defined *healthcare quality* as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" [6]. World Health Organization (WHO) defined *quality* as "the extent to which health care services provided to individuals and patients improve desired health outcomes". To achieve this, WHO added that healthcare must be safe, effective, timely, efficient, equitable and people-centred [7], specific to maternal, neonatal and child health (MNCH). Hulton and his colleagues defined *quality* of care as the "degree to which maternal health services for individuals and populations increase the likelihood of timely and appropriate treatment to achieve desired outcomes that are both consistent with current professional knowledge and uphold basic reproductive rights" [8].

Improving MNCH quality is paramount to reducing maternal and neonatal deaths and attaining sustainable development goals (SDGs), especially SDG3, which calls for better health for all [9]. In this view, World Health Organization (WHO) has elaborated a global vision where every woman and a newborn receive quality care throughout pregnancy, childbirth and the postnatal period under the umbrella of universal health coverage and quality [10]. To realize the vision, WHO has further developed a framework that, among other areas, focuses on quality assessment as one of the strategic areas to improve MNCH care quality [11]. Quality assessment is "a process of evaluating healthcare system performance against recognized quality standards" [12]. It is a necessary step in the process of improving quality. It tells whether the services provided are of high quality enough to make a difference in health and survival and how to drive improvement in quality of care [13]. In similar veins, a reliable and sustainable quality assessment mechanism is required to routinely measure and report the quality status [14]. However, the literature shows that most developing countries, such as Tanzania, fail to establish effective approaches for quality assessment in MNCH [15] [16] [17].

Facility assessment tools such as service availability and readiness assessment (SARA) [18], Service Provision Assessment (SPA) [19], Service Delivery Indicator (SDI) [20] and Needs Assessment of Emergency Obstetric have been used in assessing MNCH quality [21]. Also, approaches like demographic and health surveys (DHS) [22] [23], quality evaluation frameworks [24] [25], direct observation of clinical consultations [26], interviews with service providers and exit interviews with clients [26] and health surveys methods [27] [28] to mention a few have been commonly applied to meet various MNCH quality assessment needs. However, it has been observed that most of these quality assessment approaches currently used are resource inefficient. Most of the approaches are manual and paper-based; they require the physical presence of quality assessment personnel at the facility and may also add extra financial and human resources needs in accomplishing quality assessment [29] [30] [31]. From this point of view, inadequacy in the resources needed for quality measurement limits quality measurement and impacts the effectiveness of the approaches in quality assessment and reporting. To address these notable challenges, this study proposed and developed a machine learning-based quality assessment model that is resource-efficient for practical quality assessment in MNCH care and applicable in developing countries with resource constraints.

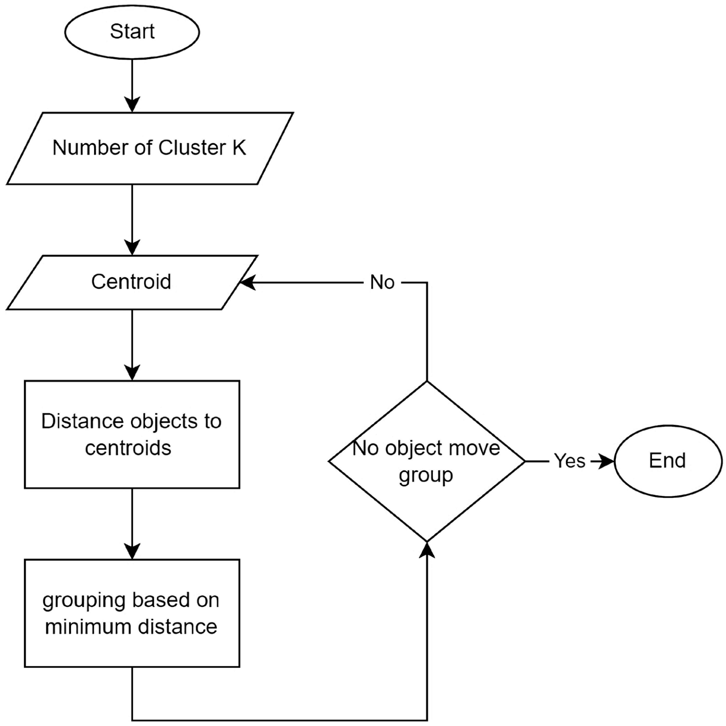
K-means is an unsupervised learning algorithm used to solve clustering problems in machine learning [31]. *K*-means clustering groups the unlabeled datasets into different clusters according to the initial number of clusters (K). *K*-means allows clustering of data into different groups; it is a convenient way to discover the categories of groups in the unlabeled datasets without needing any training [32] [33]. **Table 1** illustrates the working of *K*-means algorithm. It takes the unlabeled dataset as input, divides the dataset into *K* number of clusters, and repeats the process until it does not find the best cluster. The value of *K* is always predetermined in the *K*-means algorithm, which mainly performs two tasks; determining the best value for *K* centre centroids by an iterative process and assigning each data point to its closest *K*-center. Hence each cluster has data points with some commonalities, and it is away from other clusters [34]. **Figure 1(a)** and **Figure 1(b)** show assignment of data points into clusters.

2. Related Works

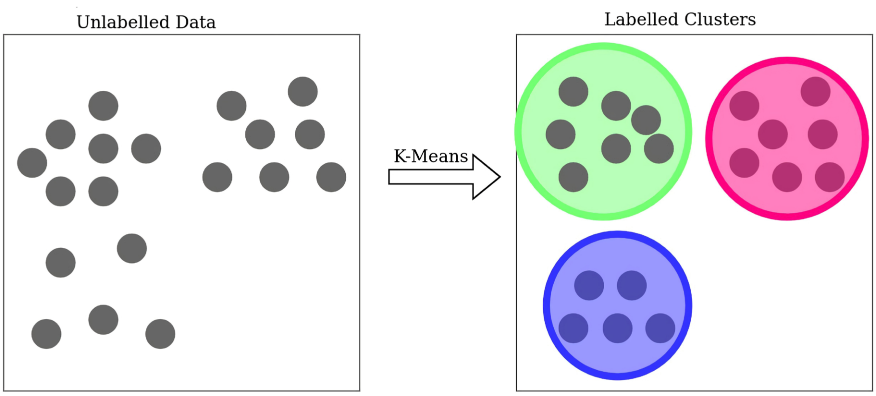
Artificial intelligence (AI) technologies, especially machine learning and data mining, are increasingly used in predicting healthcare outcomes [35] [36]. Several studies have reported to use machine learning (ML) to predict healthcare costs, utilization of resources and quality of services [37]. The ability of ML to provide insight from both structured and unstructured data makes it applicable for quality assessment in healthcare [38]. Machine learning algorithms have been used to assess the quality of healthcare services, information and treatment provided online via web pages and blogs [39]. A study by [40] used a support

Table 1. *K*-means algorithm.

The working of the *K*-means algorithm
1. Select the number of *K* (to decide the number of clusters) 2. Select random *K* points or centroids 3. Assign each data point to their closest centroid 4. Calculate the variance and place a new centroid of each cluster 5. Repeat the third step (which means reassigning each data point to the new closes centroid of each cluster) 6. If any reassignment occurs, then go to step 4; else, go to finish 7. Finish



(a)



(b)

Figure 1. (a) *K*-means algorithm Flowchart, (b) *K*-means clustering.

vector machine (SMV) algorithm to develop a model to filter and identify the web pages that do not contain quality cancer treatment information. Specifically, the model performed text categorization to identify unproven cancer treatments on the web pages. The developed model can identify web pages that make unproven cancer treatment fully automatic and substantially better than the previous web-based tools and state-of-the-art search engine technologies. The study by [41] developed a supervised learning approach using the Support Vector Machine (SVMLight) toolkit. The approach was designed to predict the reliability of medical webpages automatically. The algorithm successfully classified medical webpages as being reliable or not, based on the information web page contents and features contained, with an accuracy of 80 per cent.

Furthermore, the quality of online health information on the web pages automatically was measured based on their contents using natural language processing (NLP) and machine learning (ML) techniques [42]. Afsana and her colleagues achieved an accuracy of 84 to 90 per cent with a data mining approach developed to automatically assess the quality of online healthcare articles based on the predefined quality criteria [43]. And study by [44] developed a machine learning model using observational data to assess the quality of glycemic control care provided to patients with type II diabetes using the DIABETIMSS program. The reviewed studies have successfully assessed the quality of health care provided through the web pages. Previous studies have successfully assessed the quality of health care provided through web pages. This current study aims to use routine health data that representing care that a person has received from health care providers.

3. The Proposed Approach

The proposed quality measurement approach intends to facilitate quality measurement using *K*-means clustering. *K*-means clustering is one of the simplest and most popular unsupervised machine learning algorithms. Typically, unsupervised algorithms make inferences from datasets using only input vectors without referring to known or labelled outcomes. The proposed solution is organized into five phases. Phase I, Dataset Selection; Phase II, Dataset cleaning and preprocessing; Phase III, *K*-means clustering; Phase IV, cluster evaluation; Phase V, Cluster analysis.

K-means can describe hidden structures from unlabeled data [45]. This ability is considered as strength of the *K*-means algorithm; it is used to group the maternal and child health dataset into clusters according to their similarity and differences. Then cluster analysis was done to discover knowledge on the quality of health services from the data records in each cluster [46]. **Figure 2** illustrates the general objective of the proposed approach.

3.1. Phase I: Dataset Selection

Data selection is a process of determining the appropriate data type and source. It was employed to extract required data from District Health Information System

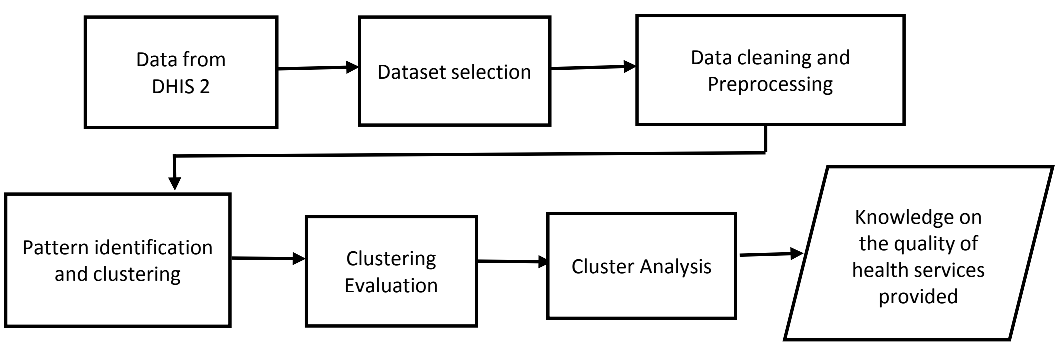


Figure 2. Proposed approach generic diagram.

(DHIS 2). DHIS 2 is currently used as a centralized database and national health information system by several developing countries, Tanzania inclusive [47]. Routinely, aggregated health data from all health centres are collected and stored in DHIS 2 [48]. MNCH data was selected to form a dataset for this study. The dataset comprised five years of maternal, neonatal and child health data (2014 to 2018) from all districts in Tanzania.

3.2. Phase II: Dataset Cleaning and Prepossessing

Data cleaning and preprocessing involve the transformation of the raw dataset into an understandable format. This is a fundamental stage in model development because it may directly affect the outcome of the expected model. Data is often incomplete and inconsistent and is more likely to contain many errors if not thoroughly checked. In this phase dataset was prepared by performing data cleaning which included removal of incomplete and inaccurate data from the dataset. Data preprocessing was done to handle inconsistent data and removing errors within the dataset. Dataset cleaning and preprocessing were accomplished using scikit-learn.

3.3. Phase III: Clustering

K-means algorithm was used for clustering. *K*-means is one of the most popular flat clustering algorithms [34]. *K*-means assign every data point to one of the *K* clusters, where *K* is the number of clusters to be formed. A cluster in *K*-means is a sphere with a centroid at its centre of gravity. To form a cluster, *K*-means minimizes the average distance of data points from the cluster centre, which is defined as the mean or centroid of the data points in a cluster. The centroid μ of the data-points in a cluster ω is computed by Equation (1)

$$\mu(\omega) = \frac{i}{\omega} \sum_{x \varepsilon \omega} \vec{x} \quad [34] \tag{1}$$

The *K*-means algorithm typically begins by selecting *K* random seeds that are assumed to represent the centre (centroid) of the *K* initial clusters. Elbow method was used to determine the *K* random seeds. The elbow curve started from 2 to 4 (see **Figure 3**); this means two to four cluster can be found in a dataset. Because we expect to determine two major groups, $K = 2$ was selected as the initial *K* value. The algorithm then calculates each data point's distance (or similarity) from all the two *K* points. These distance values assign every data point to one of the *K* clusters. A data point is assigned to a cluster closest to it, or in other words, to the cluster whose centroid has the smallest distance from the data point out of all such *K* centroids. Once all data points are assigned to one of the *K* clusters, the centroids of all the *K* clusters are recomputed. Thereafter, the process is iterated with the new centroids as a new cluster centre. This iterative process is repeated until no new cluster assignment is done. As a terminating criterion, the residual sum of squares (RSS), a squared distance of each data point vector and its cluster centroid were summed over all vectors.

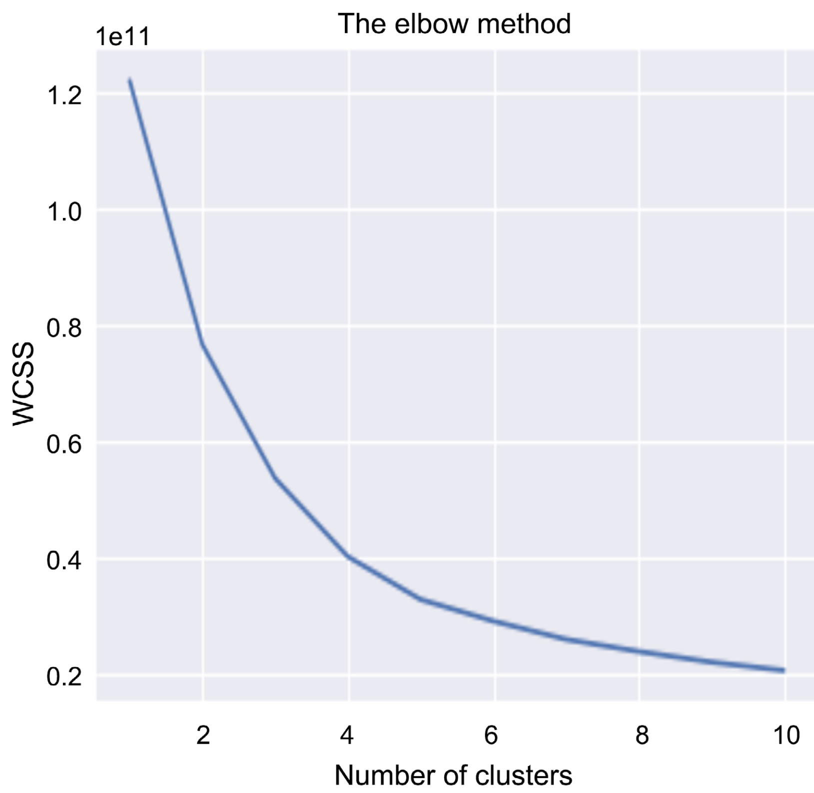


Figure 3. Elbow method.

3.4. Phase IV: Clustering Evaluation

The objective of clustering algorithms is to achieve high intracluster similarity (similar data points within a cluster) and low intercluster similarity (dissimilar data points from different clusters). This is often viewed as an internal criterion of clustering quality. The residual sum of squares (RSS), defined as the squared distance of each vector from its centroid summed over all vectors, is a measure of this internal criterion that was computed by the Equation (2) below.

$$\text{RSS} = \sum_{K=1}^{K} \text{RSS}_K \quad [34] \tag{2}$$

where

$$\text{RSS}_K = \sum_{x \epsilon \omega_K} \left| \vec{x} - \mu(\omega_K) \right|$$

A Good score of RSS (internal criterion) is a good representation of clustering quality. However, these good scores do not always turn out to be an accurate and effective measure of clustering quality. With a good RSS score, purity was deployed to evaluate the clustering quality as an alternative. Purity is among the popular external criteria for directly evaluating the clustering quality. To compute the purity, each cluster is assigned to the class which is the most frequent cluster. Then the accuracy of each assignment is measured by counting the number of correctly assigned data points, dividing by the total number of data points. Purity was then calculated by evaluating how many data points were assigned to the correct class using the formula below:-

$$\text{Purity}(\Omega, C) = \frac{1}{N}\sum_{K} \max_{j} \left|\omega_k \cap C_j\right| \quad [34] \tag{3}$$

where, $\Omega = \{w_1, w_2 \ldots w_k\}$ is the set of clusters and $C = \{c_1, c_2 \ldots c_j\}$ is the set of classes. Here w_k denotes the set of data points in the cluster w_k, and c_j denotes the data points in the class c_j. Bad clustering has purity close to "0", and perfect clustering has a purity value of "1".

3.5. Phase V: Cluster Analysis (Knowledge Discovery)

Clustering is widely used in a variety of statistical and machine learning applications. These applications can be roughly divided into two objectives. The first objective is identifying homogeneous groups within a dataset, and the second is summarizing a dataset into representative points or cluster prototypes [46]. The first objective spans a broad range of real-world problems, including the discovery of gene groups with similar functions, identifying communities in social networks and so on. In this study, clustering was performed to discover data points with similar characteristics. Similarly, the second clustering objective covers many essential problems in the current significant data era, including data summarization and comprehension. This section focuses on the latter purpose of understanding, representing a dataset by its patterns.

For easy knowledge discovery, data points in each cluster were labelled. The data points in the first cluster were labelled by "0", and those in the second cluster were labelled by "1". MNCH experts rated each data record in a cluster by giving a "good quality" score denoted by "1" and "poor quality" score denoted by "0". These labels guided knowledge discovery from each cluster. The new labels were appended to the dataset with column names "Pred" for cluster identification label and "Expert Judgment". These two columns, "Pred" and "Expert judgment", have enabled the discovery of the cluster containing data records showing high-quality and poor-quality health services. It was found that the first cluster, denoted by "0", contains data records that depict high-quality MNCH services, and the second group, denoted by "1", contains data records that depict poor-quality MNCH services. However, few data records belong in the first cluster ("0") but contain data records that depict poor quality. **Figure 4** and **Figure 5** show the dataset head and tail that illustrate the cluster analysis results.

4. Experiment Results

The model was applied to maternal, neonatal and child health dataset. The model grouped the dataset into two clusters as initialized by $K = 2$. **Figures 6(a)-(d)** present the results of four iterations conducted to obtain clustering results.

The ability of machine learning to learn and discover patterns and characteristics within the data fed onto them has enabled achievement of study objectives. The study first intended to group the data points into clusters according to their similarity and differences. Based on the study objectives; *K*-means clustering yielded best modal for this objective (**Figures 6(a)-(b)** illustrate that *K*-means

Out[38]:

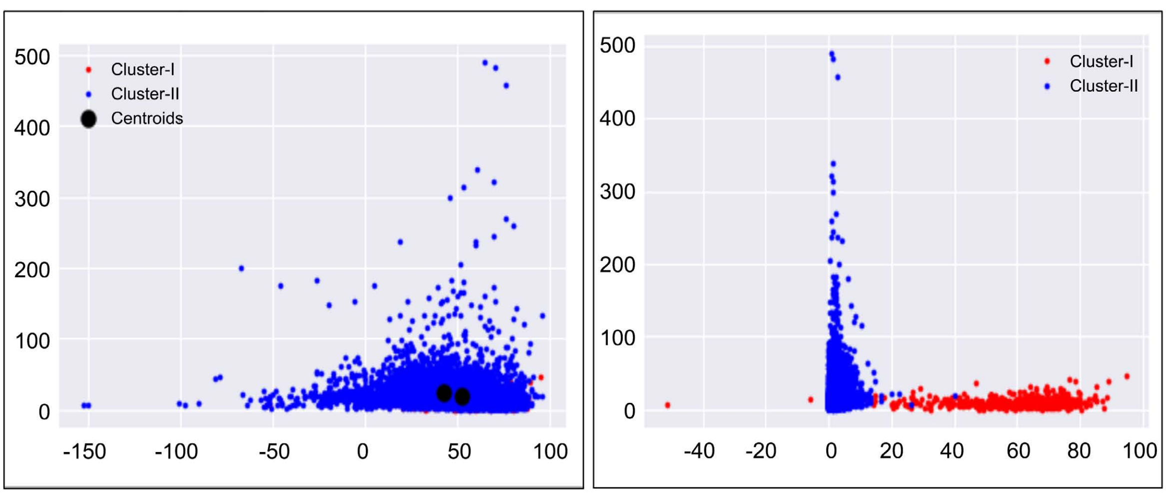
ANC HIV evalence (15-24 years)	ANC HIV prevalence 1st Test	ANC HIV prevalence 2nd Test	ANC HIV testing rate	...	Proportion of deliveries referred out	Proportion of new born alive with birth weight <2.5kg	Proportion of newborn initiated early breastfeeding	Proportion of newborn resuscitated	Proportion of women delivered at health facility	Proportion of women delivered who are less than 20 years	Still birth rate	TBA Delivery rate	pred	Expert Judgment
2.678925	1.300000	0.160000	97.500000	...	0.420000	7.500000	92.800000	2.800000	97.60000	23.200000	0.360000	0.740000	0	1
2.678925	1.600000	0.230000	97.300000	...	0.090000	3.400000	97.000000	8.100000	98.40000	20.300000	1.500000	0.110000	0	1
2.678925	2.013897	3.094317	91.245312	...	2.232117	5.802506	85.457236	9.596289	90.70771	20.395394	1.374436	1.768584	0	0
2.678925	2.013897	3.094317	91.245312	...	2.232117	5.802506	85.457236	9.596289	90.70771	20.395394	1.374436	1.768584	0	0
0.000000	3.100000	0.000000	37.800000	...	1.400000	0.000000	100.000000	10.100000	97.20000	12.700000	3.000000	1.768584	0	1

Figure 4. Dataset head.

Out[45]:

HIV nce 20 ars	ANC HIV prevalence (15-24 years)	ANC HIV prevalence 1st Test	ANC HIV prevalence 2nd Test	ANC HIV testing rate	...	Proportion of deliveries referred out	Proportion of new born alive with birth weight <2.5kg	Proportion of newborn initiated early breastfeeding	Proportion of newborn resuscitated	Proportion of women delivered at health facility	Proportion of women delivered who are less than 20 years	Still birth rate	TBA Delivery rate	pred	Expert Judgment
.85	52.678925	0.25	0.0	101.2	...	0.54	3.3	62.5	9.2	97.7	15.7	0.42	0.71	0	1
.70	52.678925	2.60	0.0	101.7	...	0.96	1.5	99.0	10.8	93.3	20.7	1.50	0.45	0	1
.00	52.678925	1.10	0.0	95.4	...	0.00	7.6	78.4	19.4	97.6	14.7	2.90	0.49	0	1
.80	52.678925	1.50	0.0	97.0	...	0.00	6.8	79.8	16.6	97.7	15.6	0.60	0.94	0	1
.76	52.678925	2.10	0.0	97.6	...	0.39	6.0	85.1	7.8	99.3	11.6	3.40	0.00	0	1

Figure 5. Dataset tail.



(a) (b)

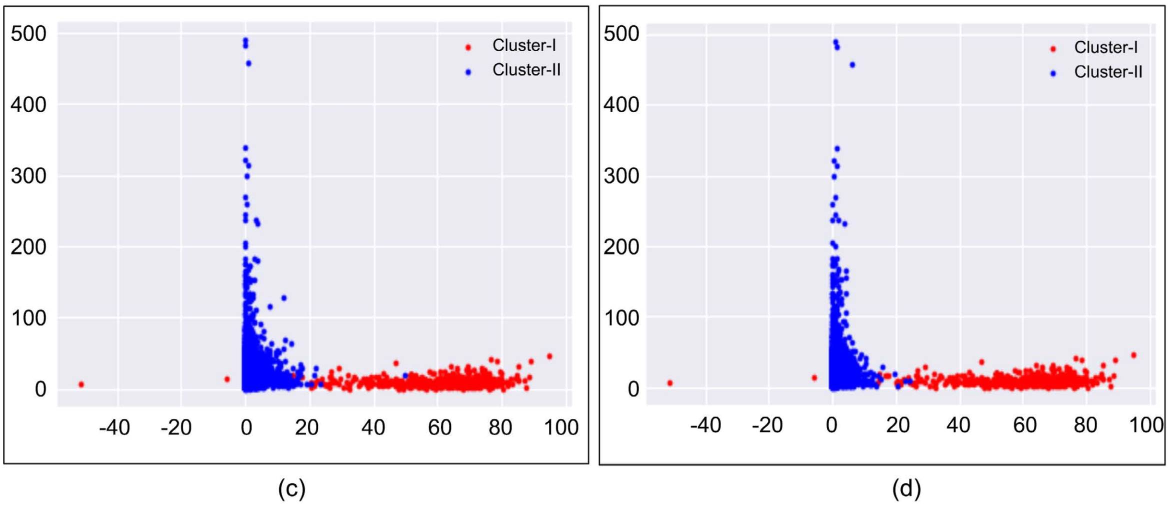


Figure 6. (a) 1st iteration, (b) 2nd iteration, (c) 3rd iteration and (d) 4th iteration.

algorithm performed well in clustering). Having very similar data points into the same group, the study also intended to discover the knowledge about health care quality depicted in each cluster. Cluster analysis showed that data records in the first cluster denoted by "0" have data records indicating the services provided are of good quality. The second cluster denoted by "1" contains data records indicating the services provided were not of good quality (Section 3.4 of this study illustrate). The results of the cluster analysis showed that the study has successful use machine learning and data from routine health services to assess the quality of MNCH services.

5. Conclusion and Future Work

This study applied the *K*-means clustering algorithm to automate the quality assessment process for MNCH services. We found that applying machine learning to assess the quality of health services provided to pregnant women and children in Tanzania is feasible. This work is not directly comparable to the existing studies because most of the existing studies assess the quality of MNCH based on healthcare structure, process and outcomes using manual and paper-based approaches and statistical analysis. This study assessed MNCH quality using routine health data from a national data warehouse. This study's results support the improvement of MNCH services quality and enable the saving of financial and human resources that traditional quality assessment approaches would need. As future work, it is intended firstly to develop and add new models for other health domains. Secondly, to develop a system that will use Google maps imagery to plot the data points from the clusters for easy results visualization and interpretation.

Acknowledgements

The authors would like to thank African Development Bank (AfDB) for funding this study.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Ayanian, J.Z. and Markel, H. (2016) Donabedian's Lasting Framework for Health Care Quality. *The New England Journal of Medicine*, **375**, 205-207. http://www.nejm.org https://doi.org/10.1056/NEJMp1605101

[2] Donabedian, A. (1988) The Quality of Care: How Can It Be Assessed? *JAMA*, **260**, 1743-1748. https://doi.org/10.1001/jama.260.12.1743

[3] Donabedian, A. (1983) Quality Assessment and Monitoring: Retrospect and Prospect. *Evaluation & the Health Professions*, **6**, 363-375. https://doi.org/10.1177/016327878300600309

[4] Nykänen, P., Brender, J., Talmon, J., De Keizer, N., Rigby, M., Beuscart-Zephir, M.C. and Ammenwerth, E. (2011) Guideline for Good Evaluation Practice in Health Informatics (GEP-HI). *International Journal of Medical Informatics*, **80**, 815-827. https://doi.org/10.1016/j.ijmedinf.2011.08.004

[5] Hulton, L., Matthews, Z., Bandali, S., Izge, A., Daroda, R. and Stones, W. (2016) Accountability for Quality of Care: Monitoring All Aspects of Quality across a Framework Adapted for Action. *International Journal of Gynecology & Obstetrics*, **132**, 110-116. https://doi.org/10.1016/j.ijgo.2015.11.005

[6] Committee on Quality of Health Care in America, Institute of Medicine (2001) Crossing the Quality Chasm: A New Health Care System for the 21st Century. National Academy Press, Washington DC.

[7] World Health Organization (2006) Quality of Care: A Process for Making Strategic Choices in Health Systems. World Health Organization, Geneva.

[8] United Nations (2016) Sustainable Development Goals Report 2016. https://unstats.un.org/sdgs/report/2016/The%20Sustainable%20Development%20Goals%20Report%202016.pdf

[9] World Health Organization (2018) Quality, Equity, Dignity: The Network to Improve Quality of Care for Maternal, Newborn and Child Health: Strategic Objectives. World Health Organization, Geneva.

[10] World Health Organization (2010) Framework for Monitoring and Evaluation of Reproductive Health Programmes in the Eastern Mediterranean Region. No. WHO-EM/WRH/081/E, World Health Organization, Geneva.

[11] World Health Organization (2016) Global Strategy for Women's. Children's and Adolescents' Health (2016-2030). World Health Organization, Geneva.

[12] Morris, C. and Bailey, K. (2014) Measuring Health Care Quality: An Overview of Quality Measures. *Families USA*, 1-16. http://familiesusa.org/sites/default/files/product_documents/HSI%20Quality%20Measurement_Brief_final_web.pdf

[13] Kruk, M.E., Leslie, H.H., Verguet, S., Mbaruku, G.M., Adanu, R.M. and Langer, A. (2016) Quality of Basic Maternal Care Functions in Health Facilities of Five African Countries: An Analysis of National Health System Surveys. *The Lancet Global Health*, **4**, e845-e855. https://doi.org/10.1016/S2214-109X(16)30180-2

[14] Mothupi, M.C., Knight, L. and Tabana, H. (2018) Measurement Approaches in Continuum of Care for Maternal Health: A Critical Interpretive Synthesis of Evidence

from LMICs and Its Implications for the South African Context. *BMC Health Services Research*, **18**, Article No. 539. https://doi.org/10.1186/s12913-018-3278-4

[15] World Health Organization (2014) Service Availability Mapping (SAM). https://www.unfpa.org/resources/service-availability-mapping-sam#

[16] Akachi, Y. and Kruk, M.E. (2017) Quality of Care: Measuring a Neglected Driver of Improved Health. *Bulletin of the World Health Organization*, **95**, 465-472. https://doi.org/10.2471/BLT.16.180190

[17] World Health Organization (2013) Service Availability and Readiness Assessment (SARA): An Annual Monitoring System for Service Delivery: Reference Manual. No. WHO/HIS/HSI/RME/2013/1, World Health Organization, Geneva. https://cdn.who.int/media/docs/default-source/service-availability-and-readinessassessment(sara)/sara_reference_manual_chapter3.pdf?sfvrsn=7a0f3a33_3

[18] Krüger, C., Heinzel-Gutenbrunner, M. and Ali, M. (2017) Adherence to the Integrated Management of Childhood Illness Guidelines in Namibia, Kenya, Tanzania and Uganda: Evidence from the National Service Provision Assessment Surveys. *BMC Health Services Research*, **17**, Article No. 822. https://doi.org/10.1186/s12913-017-2781-3

[19] World Health Organization, Organization for Economic Co-Operation and Development, and the World Bank (2018) Delivering Quality Health Services: A Global Imperative for Universal Health Coverage. World Health Organization, Geneva.

[20] Banke-Thomas, A., Wright, K., Sonoiki, O., Banke-Thomas, O., Ajayi, B., Ilozumba, O. and Akinola, O. (2016) Assessing Emergency Obstetric Care Provision in Low- and Middle-Income Countries: A Systematic Review of the Application of Global Guidelines. *Global Health Action*, **9**, Article No. 31880. https://doi.org/10.3402/gha.v9.31880

[21] Dettrick, Z., Gouda, H.N., Hodge, A. and Jimenez-Soto, E. (2016) Measuring Quality of Maternal and Newborn Care in Developing Countries Using Demographic and Health Surveys. *PLOS ONE*, **11**, Article ID: e0157110. https://doi.org/10.1371/journal.pone.0157110

[22] Benova, L., Tunçalp, Ö., Moran, A.C. and Campbell, O.M.R. (2018) Not Just a Number: Examining Coverage and Content of Antenatal Care in Low-Income and Middle-Income Countries. *BMJ Global Health*, **3**, Article ID: e000779. https://doi.org/10.1136/bmjgh-2018-000779

[23] Hulton, L., Matthews, Z. and Stones, R.W. (2000) A Framework for the Evaluation of Quality of Care in Maternity Services. University of Southampton, Southampton. http://eprints.soton.ac.uk/40965/

[24] Bhattacharyya, O., Mossman, K., Ginther, J., Hayden, L., Sohal, R., Cha, J. and Mitchell, W. (2015) Assessing Health Program Performance in Low-and Middle-income Countries: Building a Feasible, Credible, and Comprehensive Framework. *Globalization and Health*, **11**, Article No. 51. https://doi.org/10.1186/s12992-015-0137-5 http://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-015-0137-5

[25] LeFevre, A., Mpembeni, R., Kilewo, C., Yang, A., An, S., Mohan, D. and George, A.S. (2018) Program Assessment of Efforts to Improve the Quality of Postpartum Counselling in Health Centers in Morogoro Region, Tanzania. *BMC Pregnancy and Childbirth*, **18**, Article No.282. https://doi.org/10.1186/s12884-018-1906-y

[26] Duysburgh, E., Temmerman, M., Yé, M., Williams, A., Massawe, S., Williams, J. and Blank, A. (2016) Quality of Antenatal and Childbirth Care in Rural Health Facilities in Burkina Faso, Ghana and Tanzania: An Intervention Study. *Tropical Medicine &*

International Health, **21**, 70-83. https://doi.org/10.1111/tmi.12627

[27] Tripathi, V., Stanton, C., Strobino, D. and Bartlett, L. (2019) Measuring the Quality of Maternal and Care Processes at the Time of Delivery in Sub-Saharan Africa: Development and Validation of a Short Index. *BMC Pregnancy and Childbirth*, **19**, Article No. 133. https://doi.org/10.1186/s12884-019-2281-z

[28] Koblinsky, M., Moyer, C.A., Calvert, C., Campbell, J., Campbell, O.M., Feigl, A.B. and Langer, A. (2016) Quality Maternity Care for Every Woman, Everywhere: A Call to Action. *The Lancet*, **388**, 2307-2320. https://doi.org/10.1016/S0140-6736(16)31333-2

[29] World Health Organization (2014) Consultation on Improving Measurement of the Quality of Maternal, Newborn and Child Care in Health Facilities. World Health Organization, Geneva.

[30] Sprockett, A. (2017) Review of Quality Assessment Tools for Family Planning Programmes in Low-and Middle-Income Countries. *Health Policy and Planning*, **32**, 292-302.

[31] Iqbal, A., Shil a Jabed, M. and Chowdhury, M. (2022) An Improved K-means Clustering Algorithm towards an Efficient Data-Driven Modeling. *Annals of Data Science*. https://doi.org/10.1007/s40745-022-00428-2

[32] Sarmiento, A., Fondón, I., Durán-Díaz, I. and Cruces, S. (2019) Centroid-Based Clustering with αβ-Divergences. *Entropy*, **21**, Article No. 196. https://doi.org/10.3390/e21020196

[33] Sinaga, K.P., and Yang, M.S. (2020) Unsupervised K-Means Clustering Algorithm. *IEEE Access*, **8**, 80716-80727.

[34] Morissette, L. and Chartier, S. (2013) The K-Means Clustering Technique: General Considerations and Implementation in Mathematica. *Tutorials in Quantitative Methods for Psychology*, **9**, 15-24. https://doi.org/10.20982/tqmp.09.1.p015

[35] Alotaibi, S.R. (2020) Applications of Artificial Intelligence and Big Data Analytics in M-Health: A Healthcare System Perspective. *Journal of Healthcare Engineering*, **2020**, Article ID: 8894694.

[36] Ismail, A.A. M., Ali, N., Amirul, M.S., Endut, R. and Aljunid, S.A. (2021) Prediction Model for Spectroscopy Using Python Programming. *International Journal of Nanoelectronics & Materials*, **14**, Article ID: e000101.

[37] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S. and Wang, Y. (2017) Artificial Intelligence in Healthcare: Past, Present and Future. *Stroke and Vascular Neurology*, **2**.

[38] Sondhi, P., Vydiswaran, V.G. and Zhai, C. (2012) Reliability Prediction of Webpages in the Medical Domain. 34*th European Conference on Information Retrieval*, Barcelona, 1-5 April 2012, 219-231.

[39] Aphinyanaphongs, Y. and Aliferis, C. (2007, January) Text Categorization Models for Identifying Unproven Cancer Treatments on the Web. *Studies in Health Technology and Informatics*, **129**, 968-972.

[40] Al-Jefri, M.M., Evans, R., Ghezzi, P. and Uchyigit, G. (2017) Using Machine Learning for Automatic Identification of Evidence-Based Health Information on the Web. *Proceedings of the* 2017 *International Conference on Digital Health*, London, 2-5 July 2017, 167-174. https://doi.org/10.1145/3079452.3079470

[41] Afsana, F., Kabir, M.A., Hassan, N. and Paul, M. (2020) Automatically Assessing Quality of Online Health Articles. *IEEE Journal of Biomedical and Health Informatics*, **25**, 591-601. https://doi.org/10.1109/JBHI.2020.3032479

[42] You, Y., Doubova, S.V., Pinto-Masis, D., Pérez-Cuevas, R., Borja-Aburto, V.H. and

Hubbard, A. (2019) Application of Machine Learning Methodology to Assess the Performance of DIABETIMSS Program for Patients with Type 2 Diabetes in Family Medicine Clinics in Mexico. *BMC Medical Informatics and Decision Making*, **19**, Article No. 221. https://doi.org/10.1186/s12911-019-0950-5

[43] Ghahramani, Z. (2004) Unsupervised Learning. *Summer School on Machine Learning* 2003, Canberra, 2-14 February 2003, 72-112.

[44] Goroshin, R., Bruna, J., Tompson, J., Eigen, D. and LeCun, Y. (2015) Unsupervised Feature Learning from Temporal Data. ArXiv Preprint ArXiv: 1504.02518.

[45] Karuri, J., Waiganjo, P., Daniel, O. and Manya, A. (2014) DHIS2: The Tool to Improve Health Data Demand and Use in Kenya. *Journal of Health Informatics in Developing Countries*, **8**, 38-60.

[46] Dehnavieh, R., Haghdoost, A., Khosravi, A., Hoseinabadi, F., Rahimi, H., Poursheikhali, A. and Aghamohamadi, S. (2019) The District Health Information System (DHIS2): A Literature Review and Meta-Synthesis of Its Strengths and Operational Challenges Based on the Experiences of 11 Countries. *Health Information Management Journal*, **48**, 62-75. https://doi.org/10.1177/1833358318777713

[47] Bauckhage, C. (2015) Lecture Notes on Data Science: Soft K-Means Clustering. Technical Report, University of Bonn, Bonn.

[48] Khan, D.M. and Mohamudally, N. (2010) An Agent Oriented Approach for Implementation of the Range Method of Initial Centroids in K-Means Clustering Data Mining Algorithm. *International Journal of Information Processing and Management*, **1**, 104-113.