ZEGU
ZIMBABWE
EZEKIEL GUTI
UNIVERSITY
Committed to Excellence

Oikos

The Zimbabwe Ezekiel Guti University Bulletin of Ecology, Science Technology, Agriculture, Food Systems Review and Advancement

ISSN 2957-8434 (Print)
ISSN 3007-2883 (Online)

Vol. 3 Issues (1&2), 2024

Published by the Zimbabwe Ezekiel Guti University Press
Stand No. 1901 Barrassie Rd,
Off Shamva Road
Box 350, Bindura
Zimbabwe

Typeset by Divine Graphics
Printed by Divine Graphics

SUBSCRIPTION AND RATES
Zimbabwe Ezekiel Guti University Press Office
Stand No. 1901 Barrassie Rd,
Off Shamva Road
Box 350
Bindura, Zimbabwe
Telephone: ++263 8 677 006 136 | +263 779 279 912
E-mail: zegupress@admin.uz.ac.zw
http://www.zegu.ac.zw/press

About the Journal

JOURNAL PURPOSE

The purpose of *the Oikos - The Zimbabwe Ezekiel Guti University Bulletin of Ecology, Science Technology, Agriculture and Food Systems Review and Advancement* is to provide a forum for scientific and technological solutions based on a systems approach and thinking as the bedrock of intervention.

CONTRIBUTION AND READERSHIP

Natural scientists, engineering experts, technologists, multidisciplinary teams are encouraged.

JOURNAL SPECIFICATIONS

Oikos - The Zimbabwe Ezekiel Guti University Bulletin of Ecology, Science Technology, Agriculture and Food Systems Review and Advancement

ISSN 2957-8434(Print)
ISSN 3007-2883(Online)

SCOPE AND FOCUS

The journal is a forum for the discussion of ideas, scholarly opinions and case studies of natural and physical science with a high proclivity to multidisciplinary approaches. The journal is produced bi-annually.

Guidelines for Authors for the *Oikos* Journal

Articles must be original contributions, not previously published and should not be under consideration for publishing elsewhere.

Manuscript Submission: Articles submitted to the *Oikos - The Zimbabwe Ezekiel Guti University Bulletin of Ecology, Science Technology, Agriculture and Food Systems Review and Advancement* are reviewed using the double-blind peer review system. The author's name(s) must not be included in the main text or running heads and footers.

A total number of words: 5000-7000 words and set in 12-point font size with 1.5 line spacing.
Language: British/UK English
Title: must capture the gist and scope of the article
Names of authors: beginning with the first name and ending with the surname
Affiliation of authors: must be footnoted, showing the department and institution or organisation.
Abstract: must be 200 words
Keywords: must be five or six containing words that are not in the title
Body: Where there are four authors or more, use *et al.*
Italicise *et al., ibid.,* words that are not English, not names of people or organisations, etc. When using more than one citation confirming the same point, state the point and bracket them in one bracket and in ascending order of dates and alphabetically separated by semi-colon e.g. (Falkenmark, 1989, 1990; Reddy, 2002; Dagdeviren and Robertson, 2011; Jacobsen *et al.,* 2012).

Referencing Style: Please follow the Harvard referencing style in that:
— In-text, citations should state the author, date and sometimes the page numbers.
— The reference list, entered alphabetically, must include all the works cited in the article.

In the reference list, use the following guidelines, religiously:

Source from a Journal

Anim, D.O and Ofori-Asenso, R (2020). Water Scarcity and COVID-19 in Sub-Saharan Africa. *The Journal of Infection*, *81*(2), 108-09.

Banana, E, Chitekwe-Biti, B and Walnycki, A (2015). Co-Producing Inclusive City-Wide Sanitation Strategies: Lessons from Chinhoyi, Zimbabwe. *Environment and Urbanisation*, *27*(1), 35-54.

Neal, M.J. (2020). COVID-19 and Water Resources Management: Reframing Our Priorities as a Water Sector. *Water International*, *45*(5), 435-440.

Source from an Online Link

Armitage, N, Fisher-Jeffes L, Carden K, Winter K et al. *(*2014). Water Research Commission: Water-sensitive Urban Design (WSUD) for South Africa: Framework and Guidelines. Available online: https://www.greencape.co.za/assets/Water-Sector-Desk-Content/WRC-Water-sensitive-urban-design-WSUD-for-South-Africa-framework-and-guidelines-2014.pdf. Accessed on 23 July 2020.

Source from a Published Book

Max-Neef, M. (1991). Human Scale Development: Concepts, Applications and Further Reflections, London: Apex Press.

Source from a Government Department (Reports or Plans)

National Water Commission (2004). Intergovernmental Agreement on a National Water Initiative. Commonwealth of Australia and the Governments of New South Wales, Victoria, Queensland, South Australia, the Australian Capital Territory and the Northern Territory. Available online:https://www.pc.gov.au/inquiries/completed/water-reform/national-water-initiative-agreement-2004.pdf. Accessed on 27 June 2020.

The source from an online Newspaper article

Herald, The (2020). Harare City Could Have Used Lockdown to Clean Mbare Market. *The Herald*, 14 April 2020. Available online: https://www.herald.co.zw/harare-city-could-have-used-lockdown-to-clean-mbare-market/. Accessed on 24 June 2020.

Artificial Neural Network Model Building in Predicting Maternal Mortality at Chitungwiza Central Hospital, Zimbabwe

TINASHE CHIPARAUSHE[1], LOYCE GONZO[2] AND HALLELUAH CHIRISA[3]

Abstract

The study sought to apply machine learning methods to predict maternal mortality at Chitungwiza Central Hospital, Zimbabwe. The study was motivated by the need to fulfil the World Health Organization's 3.1 aspiration of the SDGs which contemplates reaching a reduced maternal mortality ratio of less than 70 per 10 000 live births by 2030. More specifically, the study sought to examine factors with statistical significance in predicting maternal mortality and adopt Artificial Neural Network classification model to effect the predictions, and evaluate the performance thereof. Secondary data were collected from secondary data base DHIs at Chitungwiza Central Hospital. Utilising a multi-layer perceptron model, it was found that causes of death, booked, mode of delivery and marital status were statistically significant in predicting maternal mortality. The accuracy of 6.0 from the confusion matrix reflects that the model is a reliable predictor of maternal mortality. In conclusion, the MLP was found to be an effective model for predicting maternal mortality depending on the metrics used on the outcome feature. The input layer consisted of 12

[1] Directorate of Research and Innovation, Zimbabwe Ezekiel Guti University, Harare, Zimbabwe, 0009-0003-8483-8790.

[2] Department of Applied Mathematics and Statistics, Midlands State University, Gweru, Zimbabwe, https://orcid.org/0000-0003-1214-0417

[3] Directorate of Research and Innovation, Zimbabwe Ezekiel Guti University, Harare, Zimbabwe; Department of Geography, University of the Free State, Bloemfontein, South Africa, https://orcid.org/0009-0003-9388-9135

nodes. The model had 2 hidden layers each with 12 neurons. The output layer had 1 node for output. All the layers were activated using the logistic function. The overarching aim of the study is to build a predictive model that will be used to predict maternal mortality.

Keywords: Sustainability, demography, policy, maternal death, management, information, data

INTRODUCTION

Sustainable Development Goals (SDGs) have identified the reduction of maternal mortality as a key priority at global level. The target is to reduce the global maternal mortality ratio to 70 deaths per 100 000 live births by 2030 and to ensure that no country with maternal mortality of more than as twice as the global average (Alemayehu, 2019). Pursuant to this target, every country is required to calculate and achieve its national target by 2030. In 2017, the World Health Organization (WHO) reported that about 810 women died every day due to preventable causes related to pregnancy and childbirth. This number remained consistent throughout the year. Between 2017 and 2 000 the maternal mortality ratio (MMR number of maternal deaths per 100 000 live births) dropped by about 38% worldwide; of which, 94% of all maternal deaths occur in low and lower middle-income countries. This provides an insight of variations in maternal mortality exist in high-income European countries (Azhar, 2018).

Women who live in low-income countries face a greater chance of dying from maternal-related causes over the course of their lives. This means that there is a higher probability that a 15-year-old woman living in such countries will die due to complications related to pregnancy and childbirth. In high-income countries this is 1 in 5300 as compared to a 1 in 49 in low-income countries. In 2020, the maternal mortality rate in Sub-Saharan Africa was 545 deaths per 100,000 live births, which is significantly higher than the rate in Australia and New

Zealand, where the rate was only 4 deaths per 100,000 live births. Sub-Saharan Africa also accounted for 70% of all maternal deaths worldwide in 2020. However, Eastern Europe and Southern Asia regions achieved the greatest reduction in maternal mortality ratio overall, with a decrease of 0.7% from a maternal mortality ratio of 38 to 11 and 0.67% from MMR of 408 to 134, respectively. The United Nations Population Fund (UNFPA) notes that, in 2008, one woman died every minute during childbirth, resulting in over half a million deaths per weeks. Box *et al*. (2015) posit that maternal mortality is a significant issue in Nigeria, and the country is among the worst in the world in terms of maternal deaths. With a population of more than 140 million, Nigeria is the most populous country in Africa and the tenth most populous country in the world (Bryman, 2016). In 2007 WHO reported that Nigeria had the second highest number of maternal deaths worldwide of approximately 59000 annually. Goodfellow, Bengio and Courville (2016) stress that the maternal mortality rate in Nigeria is 545 deaths per 100,000 live births, which is the second highest in the world, after India.

All these many recorded maternal deaths are preventable and the Government obligated to act towards the reduction of the loss of women's lives. Hair et al. (2019) aver that in 2017, there were 15 countries which were classified as either "very high alert" or "high alert" due to their status as fragile states. These countries included South Sudan, Yemen, Chad, Nigeria, and Zimbabwe. The maternal mortality ratios (MMRs) in these countries ranged from 31 in Syria to 1,150 in South Sudan. In Zimbabwe, the MMR for rural areas was reported to be 168 deaths per 100,000 live births, while the MMR for urban areas was 85 deaths per 100,000 live births. Some years ago, 8 mothers died giving birth everyday due to the economic instability that took place in Zimbabwe. Hastie, Tibshirani and Friedman (2009) aver that, "mothers are now being charged for health- related expenses and many cannot afford the basic life-saving services they need." This may result in the increased number of deaths of women giving births at their homes. Although child bearing services at Chitungwiza Central Hospital are free, the deaths of pregnant women remain a major public health problem not just for Chitungwiza as a city, but Zimbabwe as a whole.

PROBLEM STATEMENT

As a developing country, Zimbabwe is struggling to find a way to reduce maternal deaths annually, despite the fact that there has been a 38% reduction in maternal mortality ratio, which has seen the decline from 342 deaths per 100,000 live births to 211 deaths per 100,000 live births. On the other hand, the maternal mortality rate in South Sudan, according to UN inter-agency estimates, is the highest in the world with 1,150 deaths occurring for every 100,000 live births. Maternal mortality rate at Chitungwiza Central Hospital is approximately 65, 75% per 10 000 live births per year. In comparison, the global average is 211 deaths per 100,000 live births. The foregoing reflects that the cases of maternal mortality are ever increasing despite the measures put in place to control this problem. The difference in trends between Zimbabwe as a country and Chitungwiza as a city, motivated the researchers to conduct a study on maternal mortality rate using the machine learning approach called the artificial neural networks. The following were the objectives of the study:

- To identify most important variables that contribute to maternal mortality
- To build Artificial Neural Networks model
- To predict different cases of maternal deaths at Chitungwiza Central Hospital.

The research questions which informed the study are:

- Is the rate of maternal deaths continuing to increase at the hospital?

- Which type of artificial neural network is suitable for maternal mortality rate?
- What will be the future cases of maternal deaths at Chitungwiza hospital?

SIGNIFICANCE OF THE STUDY

Prediction holds currency in maternal health studies because it offers valuable information to the Ministry of Health and Child Care which enables it to devise strategies to control maternal deaths. The researchers believe that in the future there will be reduced maternal deaths and improved maternal health. When women of reproductive age experience better health outcomes, it can lead to increased household income and economic stability for families and communities.

SCOPE OF THE STUDY

The aim of this research is to develop a plan to improve the health outcomes of individuals of reproductive age, which in turn can lead to enhanced financial stability and economic well-being for families and communities. The study will analyse monthly data from 2014 to 2022, which includes a total of 874 observations. The research will be conducted using data from the District Health Information System (DHIS) that is used by the institution and other public hospitals. The data in the study comprises pregnant women who accessed Chitungwiza hospital and died during their deliveries. The patient's information includes the age, level of education, cause of death. This information would be useful to the MoHCC and other decision makers.

LIMITATION OF THE STUDY

Important information of the patients might be incomplete since the data gathered is secondary. The researcher was unable to access the data from other institutions to make a comparison between public hospitals around the country.

LITERATURE REVIEW

This section will critically examine pertinent literature in the canon. Maternal mortality rate is considered as one of the most crucial health indicators in an economy. Maternal mortality denotes the death of a woman during pregnancy, childbirth or within 42 days after the termination of pregnancy. The map below shows the distribution of maternal mortality rate around different countries.

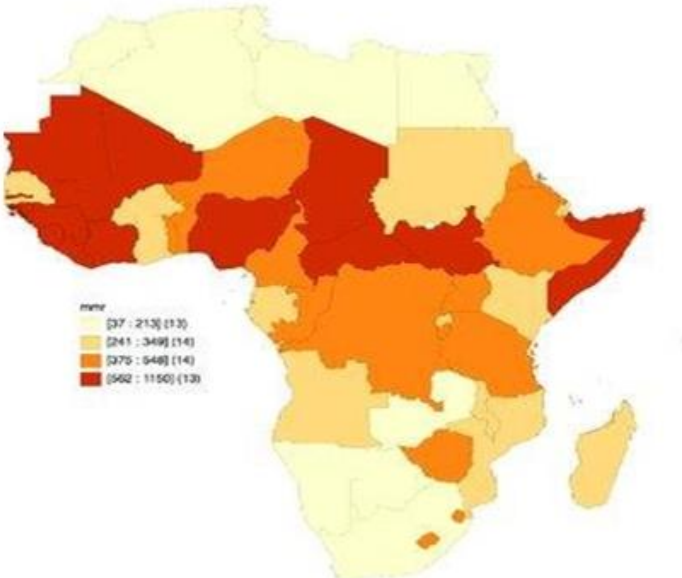

Figure 1: Maternal mortality by country 2021 (BMJ Global Health)

Despite numerous global efforts, maternal mortality continues to be a significant public health concern. Every day, around 830 women worldwide lose their lives due to complications related to pregnancy, with more than two-thirds of these deaths occurring in Africa. The researchers examined the regional distribution of maternal mortality in

Africa and explored how social determinants of health (SDoH) contribute to this distribution. The researchers collected country-level secondary data from 54 African countries, which spanned from 2006 to 2018. They obtained this data from three databases: The Global Health Observatory Data, World Development Indicator, and Human Development Report are all resources provided by the World Health Organization (WHO). The researchers conducted descriptive analyses on the data and presented their findings using tables and maps. They also utilised spatial analysis techniques, including local indicator of spatial autocorrelation maps and spatial regression. To determine and demonstrate the strength of social determinants associated with maternal mortality, the researchers used Bayesian networks, which were built in BMJ Global Health.

Their findings indicate that the average maternal mortality ratio (MMR) in African countries was 415 deaths per 100,000 live births. Spatial analyses revealed the presence of clusters or hotspots of MMR in several countries, including Guinea-Bissau, Guinea, Sierra Leone, Cote d'Ivoire, Chad, Cameroon, and Mauritania, all of which are located in the Middle and West African regions. On the contrary, the cold spot clusters were developed by two countries Namibia and South Africa from eight countries (Algeria, Tunisia, Libya, Ghana, Gabon and Congo, Equatorial Guinea and Cape Verde) which formed low–high clusters which indicates that these countries have significantly low MMR but located within the neighbourhood of countries with significantly high MMR. The results from the regression and Bayesian network analysis reveal that gender inequities and the proportion of skilled birth attendant are strongest social determinants that drive the variations in maternal mortality across Africa.

BOX-JENKINS APPROACH

The Box-Jenkins Model is a mathematical model that enables the prediction of data ranges using inputs from a defined time series

(Kibret, Chojenta and Gresham, 2015). In this study, the researchers utilised monthly data collected from KATH, Kumasi, Ghana, over a 14-year period from 2000 to 2013. They employed ARIMA models to analyse the data and determine the trend of maternal mortality. ARIMA models helped to fit the datasets. Forecasting methods are used to predict future trends or points in a population. There are two main types of forecasting methods: time series and explanatory analysis. Time series models are used to predict future patterns or trends based on past data, such as employment rates, educational attainment, incidents of crime, or disease burden in a community.

The mathematical representation of ARIMA model is as follows:
In the context of maternal mortality, the dataset Yt is said to follow an ARIMA model if the dth differences (∇dYt) can be modelled using a stationary ARMA model. The ARIMA model is built using three key parameters: p determines the AR order, d represents the number of times differencing is required to achieve stationarity, and q determines the MA order. The ARIMA model is generally represented in the form ARIMA (p, d, q), as stated by Tebbs (2018).

$$\emptyset(\mathrm{B})\ (1-\mathrm{B})^{d}\ Y\,t = \theta(\mathrm{B})e\,t \qquad (1)$$

where AR and MA characteristic operators are

$$\emptyset(\mathrm{B}) = 1 - \emptyset 1\mathrm{B} - \emptyset 1 B^{2} - \cdots - \emptyset \mathrm{p} B^{d} \qquad (2)$$

$$\theta(\mathrm{B}) = 1 - \emptyset 1\mathrm{B} - \emptyset 2 B^{2} - \cdots - \emptyset q\ B^{d} \qquad (3)$$

And

$$(1 - B)^{d} Y\,t = \nabla^{d} Y\,t \qquad (4)$$

In the given context, the autoregressive parameter f and the moving average parameter θ need to be estimated. The difference operator (∇) and the backward shift operator (B) are also used in the estimation process. Additionally, the random process et has a mean of 0 and a variance that does not depend on time (^{2}). The researchers divided their dataset into two parts: a training sample and a testing sample. The training data comprised around 85.7% of the dataset, covering the period from 2000 to 2011, and was used to build the model. The testing sample, which contained 14.3% of the data from 2012 to 2013, was used to evaluate the validity of the model. From the estimates of RMSE, MAE and MAPE for training model and testing model respectively, it is shown that the training model had a good predictive power since estimates from both models were close. The Training model was used to predict maternal mortality because of the closeness in terms of errors in models. The results show the predicted maternal mortality cases which lie within the 95% confidence intervals. Mortality cases were seen to be constant over the time period from 2012 into 2014 for the hospital.

CONTRIBUTION OF EACH VARIABLE IN PREDICTING MATERNAL MORTALITY

In selected the districts of Nigeria, they use the descriptive survey method to predict maternal mortality according to the contribution of each variable. It was a study of all females who were at their reproductive age and married, regardless of whether the death occurred during labour or puerperium. The researchers utilised a simple random sampling technique to select their participants, choosing two from urban areas to form the first cluster and two from rural areas to form the second cluster. To ensure the reliability of their questionnaires, they pre-tested them by administering them to 200 adults of reproductive age from both urban and rural areas. The data collected was used to estimate the reliability of the instrument, using Cronbach's coefficient (R), which was found to be 0.82.

The researchers applied multiple regression statistics to test their hypothesis at a 0.005 level of significance. They note that half a million women die each year due to pregnancy-related complications, and that 95% of these deaths occur in developing countries.

The researchers found that early marriages/early childbearing had the highest standard regression weight among the variables studied, with a coefficient of determination of 0.200, F value of 401.04, and P-value of 0.001. Educational attainment was the second most important variable, with an R^2 value of 0.058, F value of 126.32, and P-value of 0.001. These findings were based on data collected from both urban and rural areas.

RANDOM FOREST

The random forest technique involves constructing multiple decision trees and combining their predictions to arrive at the final prediction. Little & Rubin (2019) reveals that programme data from women enrolled in the Safer Deliveries programme of health workers in Zanzibar was used to develop a model that accurately predicted maternal mortality and was able to identify whether a pregnant woman would deliver in a health facility or not. During the enrolment visit, data such as health characteristics and demographic information were collected. Four machine learning methods, including logistic regression, artificial neural networks, Least Absolute Shrinkage and Selection Operator (LASSO) regularized logistic regression, and additional sampling techniques, were used to address the imbalance in the data, such as adding synthetic home deliveries using SMOTE. To evaluate the model, the area under the receiver operating operator (ROC curve) was used, and the chi-square statistic was calculated to estimate the predictors that best discriminated between home and facility deliveries. The methods were applied to predict maternal deaths, and the researchers found that they correctly predicted the delivery location for 68% to 77% of the women in the test set. The model had higher accuracy in predicting facility delivery compared to delivery at home, with the random forest technique identifying 74.4% of women who delivered at home. The researchers identified education, parity, and age as important predictors, and found that community-level variables were highly predictive, indicating their importance in the prediction models.

ARTIFICIAL NEURAL NETWORK AND REGRESSION MODEL

Neural network and regression are similar in their nature and application (Makworo, 2019). They have both input and output nodes, use connection weights and bias weights and use methods like cross-entropy to train the model, however neural network requires large amounts of data to use inductive learning models. They grouped data into two sets for cross-validation that is training set which is used to develop a model and testing set that is used to evaluate the model's performance. Appropriate data splitting is a method commonly used in machine learning to reduce poor generalization that are referred to as over-training or over-fitting of models. Using more training data improves the classification and the predictive power. Munjanja and Lindmark (2013) rely on ARIMA models in analysing maternal death. They used residual plots to forecast the maternal mortality. From the graphs that he plotted he managed to find the trend that describe the increase in the rate of maternal mortality. This study is in accord with the statement made by Munjanja and Lindmark (2013) that maternal deaths remain a major challenge to the healthcare systems in Zimbabwe.

Tura & Fantahun (2019) employed multivariate analysis to model the relationship between maternal mortality and its determinants in Ethiopia, shedding light on the complex interplay of factors influencing maternal health outcome. They used the data from the World Health Organization (WHO) and United Nations Population Division to compute maternal mortality ratios and able to get the most important variables of MM. They made use of multiple regression analysis to show the relationship between MMRs and several explanatory variables such as health systems, socio economic and demographic factors. Maternal education, fertility rates and access to skilled birth attendants were most significant factors.

From their research they found out that women who were not educated has a high degree in maternal mortality as compared to those who possess better education, nations with high fertility rates were found to have a significant high number of maternal mortality than countries with low fertility rates and women who were helped by a skilled birth attendant to give birth had a lower risk of maternal mortality than those who did not.

METHODOLOGY

The aim of this section is to present meaningful results that were obtained from the data gathered using various research methods and to describe how the study was conducted. Kumar (2019) provided a comprehensive overview of research methodology, emphasizing the importance of careful planning and design. It also helps readers understand the approach and methods used to draw conclusions.

Chitungwiza Central Hospital has reproductive health coordinator who provides advice and support for all types of maternity cases and infant care. The midwifery services also provide specialist support of specific antenatal or post-natal problems. The target population in this research will be pregnant women in the wards and those who already gave birth but still in the post- natal wards.

A research design is a plan that enables one to acquire and evaluate data from the noticed facts and events. The study made use of the data that was collected from maternal departments of CCH and then compiled at national level for maternal mortality in Zimbabwe through the District Health Information System (DHIS). This is quantitative descriptive research, through gathering and statistically analysing data the researchers have used Artificial Neural Network model to predict maternal deaths to analyse the trend of deaths since 2014 till future years. Kibret, Chojenta and Gresham (2015), Azhar *et al*. (2018) and Alamayehu *et al*. (2019) applied ARIMA models to analyse and forecast MMR, highlighting the potential of this approach for public health research whilst, the researchers in this study had chosen to use a multilayer perceptron type of Artificial Neural Network (ANN) technique instead. This is because the ANN technique is better suited for handling big data.

Secondary data are information that has been previously collected and used for other purposes by Bryman (2016). In the present research study, the researchers utilised secondary data, which was obtained from the Health Information Department of the company. The data were extracted from the labour and maternal departments and was stored in an online database called the District Health Information System (DHIS), which is used by CCH and other public hospitals. The use of secondary data is cheap because it cut costs of visiting the health institution daily to collect data. The availability of this data made it a good data set as it meets characteristics such as sufficiency, accuracy, reliability, relevance and the objectives.

The target population can be defined as a group of individuals or objects. In this study the target population is the total number of maternal deaths cases recorded at CCH in Zimbabwe. The research covers a period of January 2014 to December 2022, utilizing monthly data that yields a total of 874 observations. In Zimbabwe the MMR decreased from 494 per 100 000 live births in 2014 (2.5% decline from 2013) to 363 per 100 000 live births in 2022. However, at Chitungwiza

Hospital, maternal deaths are increasing highlighting a need for targeted interventions.

The process of extracting useful information from data to aid decision-making is known as data analysis. Numerous methods and techniques can be utilised to carry out data analysis. In the current study, the researchers obtained data from the institution's database and employed various data management strategies, including data cleaning, data transformation, and data visualisation, to minimize any potential bias in data. The researchers implemented these strategies to subject the data for analysis. Both quality and integrity of the research are based on the early planning. The next chapter also depend on the data analysis plan.

SOFTWARES

Microsoft excel, SPSS and R programming are soft wares that are used for this study. The researchers chose R Programming for performing many analyses because: R language offers multiple machines learning operations, including classification and regression, which makes it an ideal choice for developing artificial neural networks. The ggplot2 library in R simplifies the process of producing high-quality plots and graphs. Compared to other programming languages, R is widely used for developing statistical tools. With over 10,000 libraries, R provides a vast array of options for users. Furthermore, R is an open-source language that can be used without the need for licensing or fees

DATA PRE-PROCESSING

Data pre-processing is a crucial step in the process of data mining and analysis, wherein raw data are transformed into a more easily understandable format. This conversion of data is necessary to enable effective processing of the data in data mining. Various tools and methods, such as sampling, transformation, imputation, normalization, and feature extraction, are employed for pre-processing data. In the case of maternal data, data may have missing information, errors due to manual input, or redundant entries. While humans can detect and correct these issues, data used for training machine learning algorithms must undergo automatic pre-processing. The primary objective of the model is to generate accurate and precise predictions, which requires the algorithm to interpret the data's features.

MISSING VALUES

Missing data or missing values refer to the lack of a data value for a variable in an observation or the absence of data that was not stored for some variables in a dataset. Part 1 reviews conform rather nicely to a nonstatistician's caricature of statistics" (Little & Rubin, 2019, p150). Furthermore, missing data can introduce bias in the estimation of parameters. Missing values are frequently found in datasets and can have a notable impact on the accuracy of models and calculations that use the data. There are several reasons why data might be missing, including instances where a value is lost, forgotten, or not properly stored, or where the value of an observation simply does not exist.

Understanding the origin of missing values is essential for determining the best approach for handling the missing data.

SPLITTING THE DATA

After preparing the data, the researcher split. The data will be divided into training and testing datasets, with a ratio of 80% for training and 20% for testing. The training set will be utilised to train the MLP model, while the testing set will be used to assess its performance.

MODEL BUILDING

ARTIFICIAL NEURAL NETWORK

According to Goodfellow, Bengio, and Courville (2016), Artificial Neural Networks are a fundamental component of deep learning, enabling machines to learn complex patterns and representations from data. It being a type of supervised machine learning that is composed of a series of nodes, also known as neurons, which are connected in parallel. Each artificial neural network has multiple interconnected nodes that are arranged in layers. Generally, ANNs can be divided into three layers of neurons. The first layer, called the input layer, receives data, while the second layer, known as the artificial neuron or hidden layer, is responsible for identifying patterns and performing internal processes. Finally, the output layer generates and presents the network's results. Neurons in one layer transmit signals to other neurons in the next layer, and this is how the network performs computations. These signals are represented as real numbers, and the output of each neuron is determined by summing its inputs, weighting them individually, adding them together, and passing the sum through a non-linear function to obtain the output. Neurons and edges have a weight, and this weight affects the signal's strength. The signal is sent when the aggregated signal crosses over a threshold that is set.

PERCEPTRON MODEL ARCHITECTURE

Rosenblatt (1958). The perceptron a probabilistic mode for information storage and organization in the brain. Psychological Review, 65(6), 386-408. It employs a different artificial neuron known as the threshold logic unit (TLU) or linear threshold unit (LTU). The inputs and output are numerical instead of binary on/off values, and each input connection is assigned a weight. The TLU calculates a weighted sum of its inputs (z = w x + w x + ⋯ + w x = x w), applies a step function to that sum, and outputs the result: h (x) = step (z), where z = x w. This algorithm enables neurons to learn and process elements in the training set one at a time. The perceptron is composed of four parts: input values, weights and bias, a net input function (weighted sum), and an activation function. The researcher utilised three layers, comprising an input layer with 12 neurons (one for each variable), two hidden layers, and a single output layer with one neuron for each variable. This can be illustrated by the following diagram:

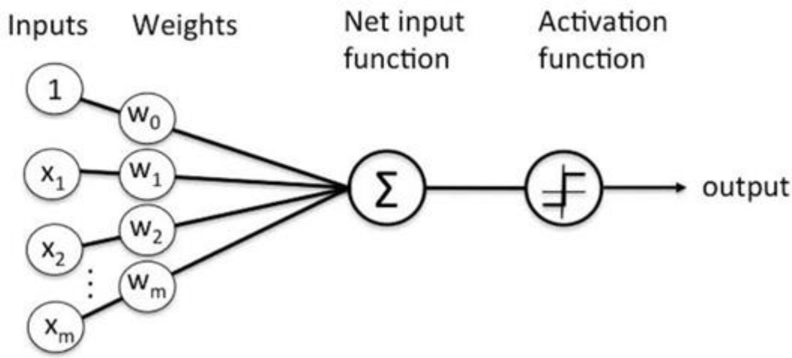

Figure 1: Architecture of the ANN model

PERCEPTRON LEARNING RULE

In a Feed Forward artificial neural network like the perceptron, the algorithm automatically learns the optimal weight coefficients. Multiple input signals are received by the perceptron, and if the sum of input signals exceeds a certain threshold, it sends a signal or does not produce an output. This rule can be utilised to forecast the class of a sample. To learn the weights for the input signals, the perceptron algorithm draws a linear decision boundary. If the desired range for outputs is between 0 to 1, the algorithm can be adjusted accordingly. This can be done by using an activation function. An activation function is a function that converts the input given (the input, in this case would be weight sum) into a certain output based on a set of rules. The original perceptron was designed to take several binary inputs and produce one binary output (0 or 1). Each output is important so the best way to represent them is to use different weights before deciding like true or false (0 or 1) respectively, the sum of the values should be greater than a threshold value.

EXPLANATORY VARIABLES

The researchers will pay attention to the following variables namely dependent and independent variables:

- Age ~ age of mother at birth

- Parity ~ number of deliveries
- Cause of death ~ factors contributing to maternal deaths
- Booked ~ whether a patient is booked or not 0 (No) and 1 (Yes)
- Marital status ~ marriage status of a patient, coded as 1= married, 2= divorced and 3= single
- Mode of delivery ~ either normal delivery (NVD) or Caesar(C/S)
- Delivered ~ patient gave birth or not 1(YES) and (NO)2
- Level of education ~ primary, secondary, higher, and non
- Occupation ~ employed or unemployed 1 and 2, respectively.
- HIV status ~ negative or positive
- Period of deaths ~ intrapartum, puerperium and antenatal.

ARTIFICIAL NEURAL NETWORK MODEL ASSUMPTIONS

ANN assumes that the input variables are independent from each other;
The relationship between the input and output variables is a linear one;
The input and output variables are normally distributed;
ANN assumes that the statistical properties of data remain constant over time or space;
ANNs require randomness in the initialization of weights and bias in the network;
It also assumes that the data are free of measurement errors and noise;

TRAINING THE MODEL

The data are trained using the training data, this process entails inputting the data into the network, calculating the output, comparing it to the desired output, and modifying the weights of the connections between neurons to decrease the error

MODEL EVALUATION

To achieve model's accuracy and performance of the predictive model researchers utilised the confusion matrix.

CONFUSION MATRIX

In the context of the confusion matrix, the rows denote the target classes present in the dataset and the output classes from the neural network that correspond to them. This is also applicable to binary classification, where the decision threshold is typically set to 0.5. Outcomes above 0.5 are taken as positive and those below that value are considered negative. The confusion matrix is illustrated in the following table:

Table 1: Structure of a confusion matrix

	Predicted positive	**Predicted negative**
Positive	TP	FN
Negative	FP	TN

True Positives (TP) ~ positive accurate classifications False Positives (FP) ~ false positive predictions
False Negatives (FN) ~ false negative predictions
True Negatives (TN) ~ negative accurate classifications

DATA ANALYSIS AND PRESENTATION

The section presents the findings of the study on the extent of maternal mortality rate at CCH, the factors that are statistically significant in predicting maternal mortality and the development of ANN's model and evaluating its effectiveness. This follows a discussion and analysis of the study's nature, literature review, and research methods. The data will be displayed using graphs, tables, and narratives to illustrate how it is distributed. As a result, the section will offer in-depth details on the visualisation, interpretation, and analysis of the research findings.

MISSING VALUES

Table 2 shows missing values that are highlighted as TRUE. False is representing values that exist in the data set, rows 14 and 19 have some missing values

Table 2: Missing and not missing values

Rows	Missing And Not Missing Values					
12	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
13	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

14	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE
15	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
16	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
17	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
18	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
19	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE
20	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
21	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE

HANDLING MISSING VALUES

The researchers used the method of listwise deletion which is simple to implement in handling missing values. This method removes completely everywhere where there appear to be missing values. This is shown in Table 3, whereby both rows 14 and 19 are removed from the data set and returns a new data frame that is cleaned with new observations. This was done to have unbiasedness, efficiency and accurate model which can result in good predictions or classifications.

Table 3: A sample of cleaned data frame

Rows	Dates	Booked	Age	Parity	Delivered	Mode	Period of deaths
12	15/03/14	Yes	19	0	No	C/S	An15tenatal
13	18/03/14	No	40	2	Yes	C/S	Puerperium
15	21/03/14	Yes	45	3	Yes	C/S	Intrapartum
16	22/03/14	Yes	45	3	No	C/S	Antenatal
17	02/04/14	Yes	34	2	Yes	NVD	Antenatal
18	07/04/14	Yes	45	2	Yes	NVD	Antenatal
20	13/04/14	Yes	35	1	Yes	NVD	Puerperium

EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis is an approach used in data analysis to summarise the main characteristics of a dataset, often with visual methods. It is an important step in understanding the dataset and uncovering patterns, trends, relationships and anomalies that may exist in the data. It can also help to identify errors, missing values, and outliers in the data, and inform the selection of appropriate statistical models for prediction or inference.

SUMMARY STATISTICS

This subsection provides the descriptive information of different features of interest which include independent variables such as age, parity, cause of deaths, marital status, level of education, cause of deaths and looks at the predictor variable Maternal Mortality. The table of summary statistics about the variables is given (Table 4)

Table 4: Summary statistics

Descriptive Statistics	Maternal Mortality	Age	Parity	Level of Education	Marital Status	Cause of deaths
Max	1	45	4	---	2	----
Min	0	16	0	---	0	----
Mean	----	30.74	1.438	---	----	----
Mode	1	31	0	---	0	----
1st Q (25%)	0	23	1	---	0	----
2nd Q (50%)	----	31	2	---	1	----
3rd Q (75%)	1	39	2	---	2	---

From the summary statistics above, there are missing values which shows that there is presence of categorical variables and they failed to provide useful numerical statistics.

INTERPRETATIONS OF THE UNIVARIATE RESULTS

MATERNAL MORTALITY

The maternal mortality variable, serving as the response variable, exhibits a binary distribution. The values range from 0, indicating death, to 1, indication survival. This analysis was carried out to identify errors between the values other than 1 and 0. The variable could not provide mean statistics because they are non-numeric variables that represent different categories or groups. The mode appeared to be 1 and it shows us the value that appears the most in that variable in the data set. This means that the institution has many survivors as compared to those who have died which are represented by the value 0. According to Figure 2, 25% were discovered to be dead, on contrary 75% were said to be survivors. This is depicted by the bar graph.

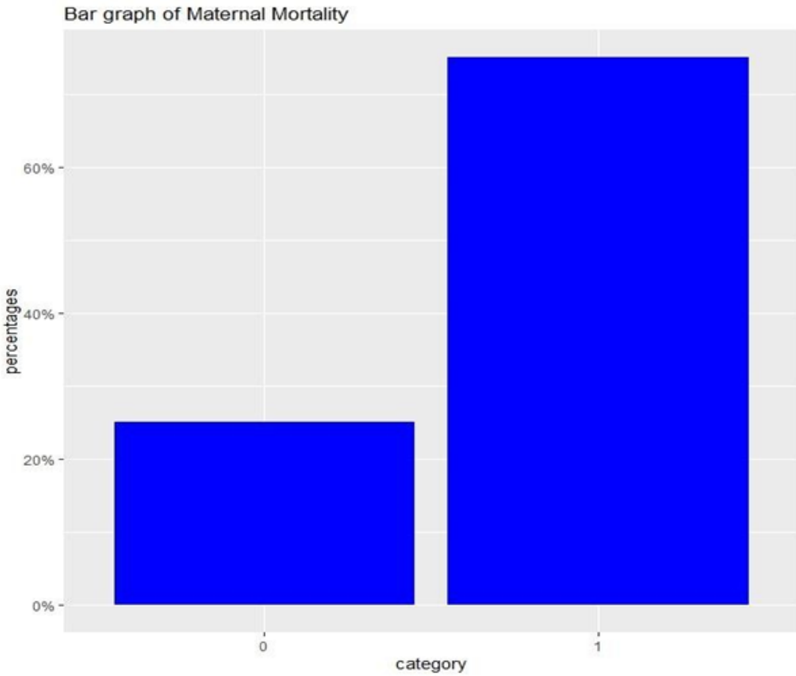

Figure 2: Univariate analysis of maternal mortality

AGE

The patient ages at the hospital range from 18 to 45 years old, with an average age of 30.74. The age that appears most frequently in maternal mortality cases is 31. Interestingly, the age distribution is balanced, with an equal number of patients above and below 31 years old.

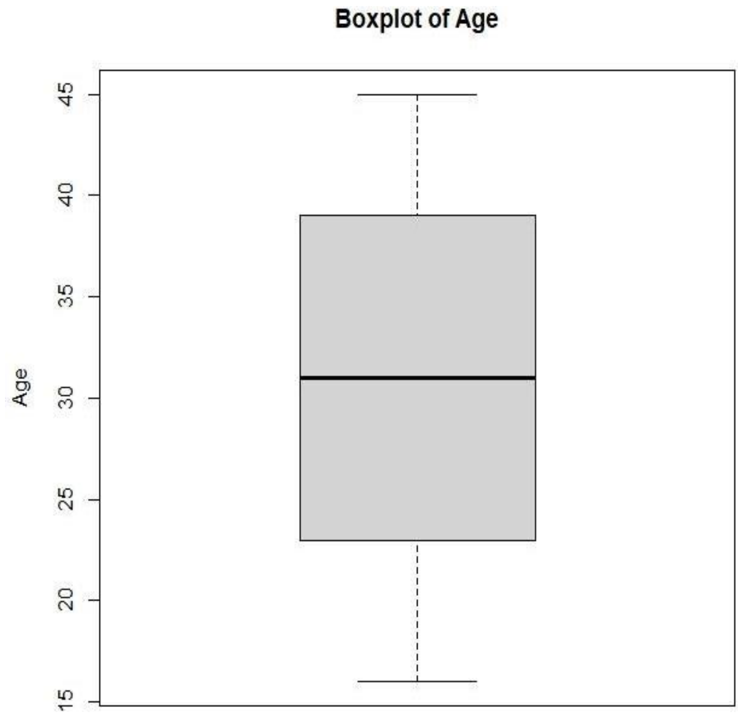

Figure 3: Univariate analysis of age

CAUSE OF DEATH

The diagram below reveals that crypto meningitis is the leading cause of maternal deaths, accounting for the highest percentage of fatalities. Placenta abruption which occurs when the placenta separates from the womb before childbirth is another significant contributor to maternal mortality. Overall, various factors are contributing to high maternal mortality rates with minimal differences in their impact.

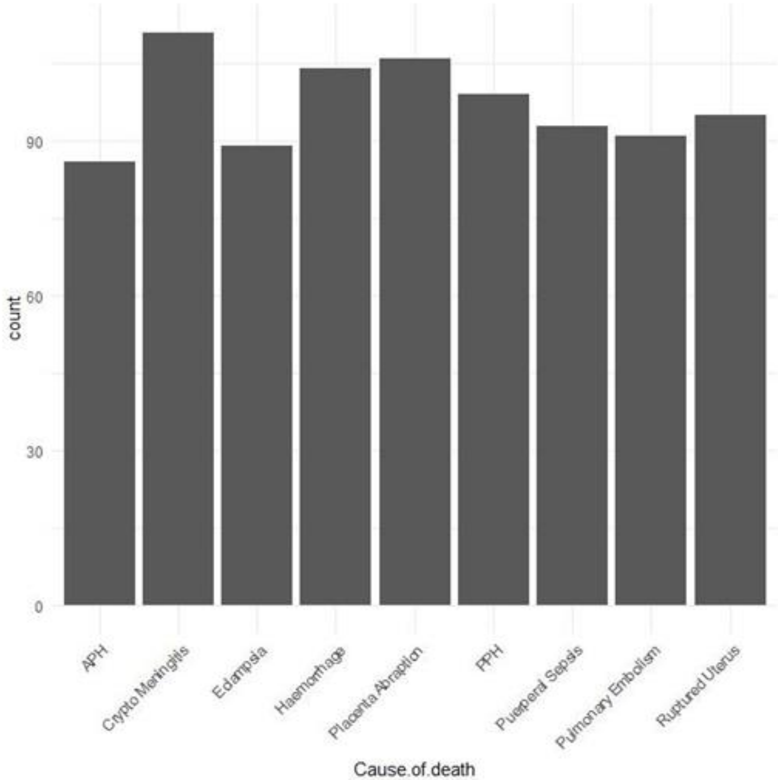

Figure 4: Univariate analysis of cause of death

MARITAL STATUS

Maternal mortality can be influenced by various factors including the mother's health, healthcare quality and socioeconomic status. Research suggests that unmarried women including singles and divorces face higher risk of maternal mortality compared to married women. This disparity may be attributed to limited healthcare access, reduced social support, and increased poverty rates. The data reviews that divorced women (44.44%) have the highest percentage, followed by singles (33.33%) and unmarried women (22.22%).

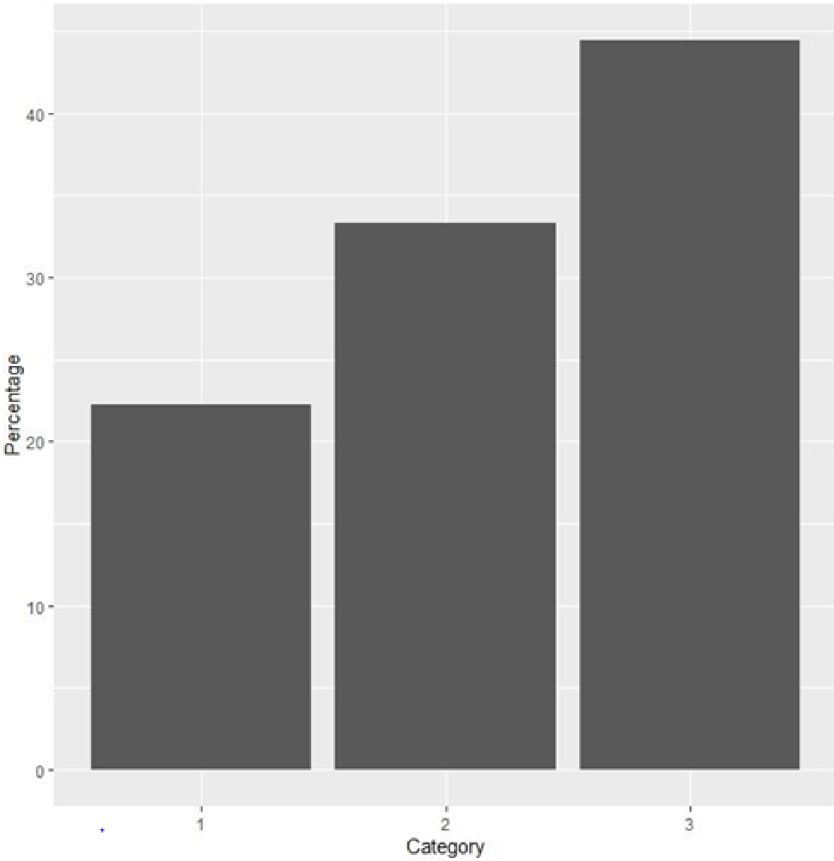

Figure 5: Univariate analysis of marital status

BIVARIATE ANALYSIS

It is a statistical method used to explore the relationship between two variables. This section applies bivariate analysis to investigate the correlation between previously examined variables and maternal mortality, aiming to determine the strength and direction of these relationships.

INDEPENDENT VARIABLES VERSUS RESPONSE VARIABLE

i.

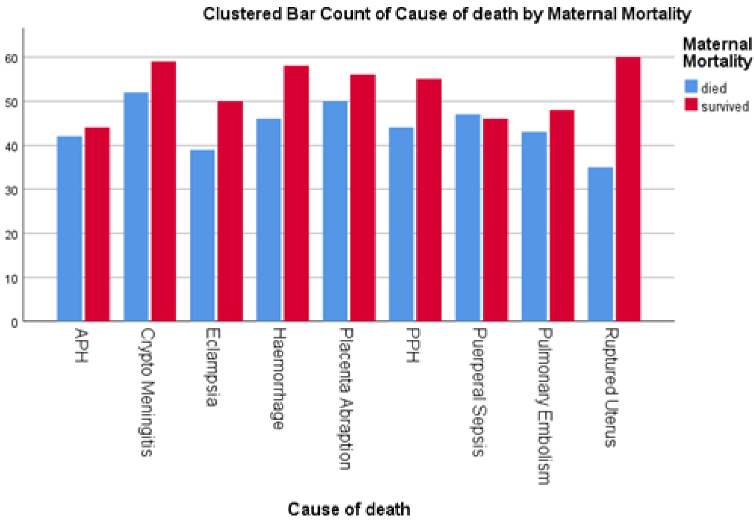

Figure 6: Bivariate analysis between maternal mortality and cause of death

Figure 6 reflects that crypto meningitis is the leading cause of death at the hospital with 52 fatalities and 59 survivors. Placenta abruption is the second-highest contributor to deaths with approximately 56 survivors. Notably the data suggests that patients with placenta abruption have a higher risk of death.

Puerperal sepsis, also known as postpartum sepsis, is a potentially life-threatening bacterial infection that occurs in women after giving birth delivery Alamaheyu *et al*. (2018). Puerperal sepsis and Antepartum Haemorrhage (APH) have a similar impact on maternal mortality with a relatively balanced number of deaths and survivors below 50. This suggests that these two conditions are contributors to maternal mortality. While puerperal sepsis has a slightly higher rate, both variables are crucial to the model, indicating their importance in understanding maternal mortality. Antepartum Haemorrhage (APH) refers to bleeding that occurs from or within the genetical tract especially between 24 weeks of pregnancy and the time of childbirth (Bryman, 2016).

Ruptured uterus is the least common cause of maternal deaths, but it is still a significant factor in the analysis. Notably the high number of survivors (around 60) and relatively low deaths can be attributed to intensive care and close monitoring of high-risk women.

ii.

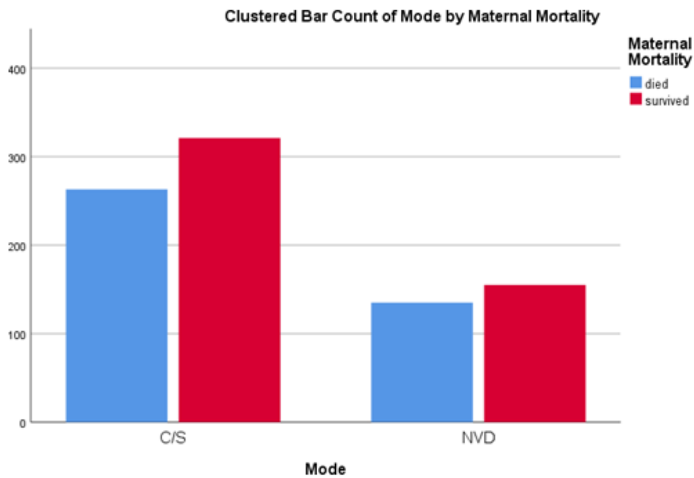

Figure 7: Bivariate analysis between mode and maternal mortality

The figure 7shows that Caesarean sections have a risk of complications compared to normal deliveries. This is likely due to the surgical nature of the procedure which requires extra caution and care unlike normal labour wards.

From the figure below shows that maternal mortality rates are significantly higher among women who were not booked (no) for prenatal care compared to those who were (yes). This disparity may be due to healthcare providers prioritizing booked patients potentially delaying critical care of adverse outcomes including death.

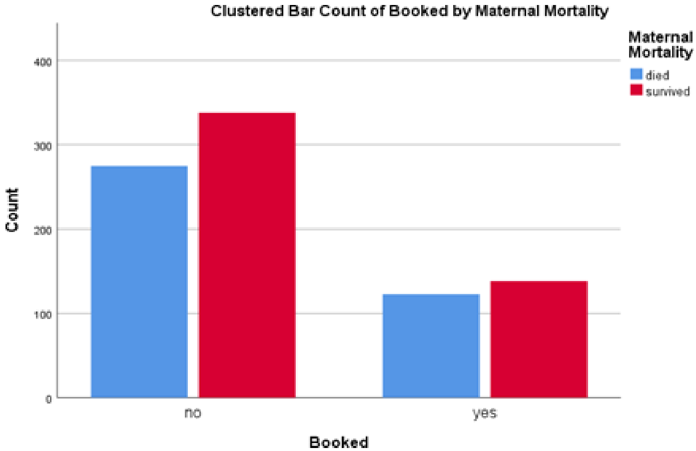

Figure 8: Bivariate analysis between booked and maternal mortality

MODEL BUILDING

MODEL FITTING

Options were entropy fitting with 29 weights decay = 0.05

Table 5: Summary of MLP

Input Layer	**Hidden Layer**		**Output Layer b-> O->0.30**	
	H1 with weights	**H2 with weights**	**H1**	**H2**
b ->	-1.93	- 1.21	4.63	-2.71
1 ->	0.00	1.28		
2 ->	-2.31	- 3.45		
3->	0.34	- 0.34		
4->	0.89	- 0.22		
5 ->	-1.29	-1.11		
6 ->	-0.50	- 1.79		
7 ->	-0.79	- 1.88		
8 ->	-0.33	- 0.26		
9 ->	2.69	3.76		
10 ->	0.04	- 1.78		
11 ->	1.29	1.97		
12 ->	0.14	1.16		

INTERPRETATION OF THE MODEL SUMMARY

The Multi-Layer Perceptron (MLP) model reveals that the neural network architecture consists of three layers: input, hidden and output. Each input and hidden layer has 12 neurons, matching the number of variables, while the output layer has 2 neurons. The model utilizes a logistic activation function and specifies parameters such as weights and decay.

CHOOSING SIGNIFICANT VARIABLES

The MLP model was trained using the dataset's attributes. The model's results below show the importance of each input variable in predicting accuracy ranked from least to most important. The top significant variables are mode of delivery (mode), cause of death, book, and marital status. Researchers used a threshold value of Pr > 0.55 to select significant variables considering variables with importance percentages above 53% as a significant.

Table 6: Choosing significant variables

Variables	Importance
Mode	100.000
Cause of death	99.347
Booked	96.566
Marital status	94.214
Occupation	52.034
Age	35.312
Level of Education	21.594
HIV Status	19.784
Parity	8.025
Period of death	5.730
Delivered	0.000

The diagram below illustrates the importance of each explanatory variable in the model. Mode of delivery stands out as the most significant variable (100%), emphasizing the need to consider the type of delivery (normal or caesarean). Cause of death follows closely (99.347%), encompassing factors like haemorrhage, sepsis, and eclampsia. Booked status (96.566%) and marital status (94.21%) are also crucial. Variables with importance below the 0.55 threshold were removed to refine the model's predictive accuracy.

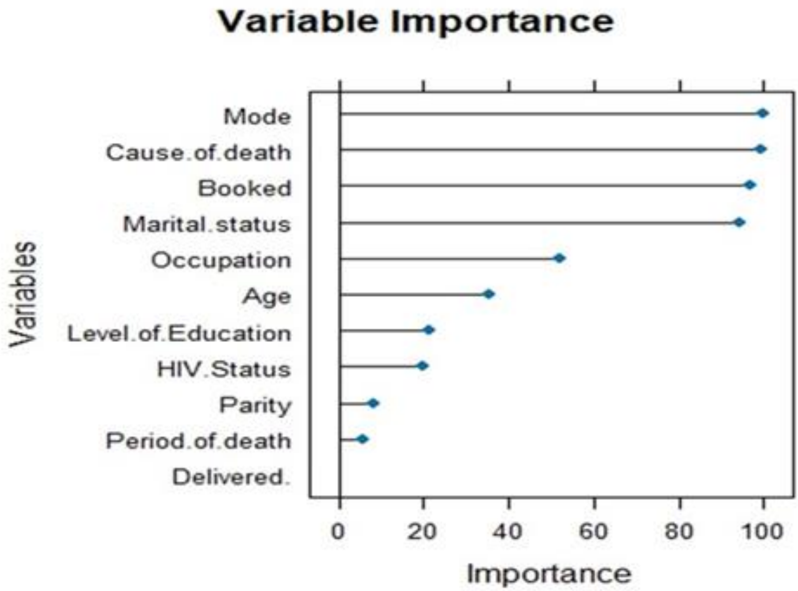

Figure 9: Variable importance

FINAL FITTED MODEL

Maternal mortality is the dependent variable with predictor variables that are significant namely mode, cause of death, booked and marital status

MODEL EVALUATION

According to Hair *et al*. (2019), a confusion matrix is a valuable tool for assessing classification model performance by comparing predicted and actual outcomes. It provides a clear visual representation of the model's errors, making it easy to understand. Little & Rubin (2019) suggest that evaluating a model engages the comparison of outcomes with objectives. The model was evaluated using the confusion matrix as shown below:

Table 7: Confusion matrix

N = 175		Predicted	
		Died	Survived
Actual	Died	71 (41% classified Correctly)	30 (17% misclassified)
	Survived	18 (10% misclassified)	56 (32% classified correctly)

The confusion matrix indicates that the MLP model is performing well as evidenced by the high values of True Negatives (TN) and True Positives (TP). TN shows correct predictions of "no" outcomes, while TP shows correct predictions of "yes" outcomes, demonstrating the model's accuracy.

The confusion matrix reveals the following predictions:

- 71 correct predictions of survival (True Negatives)
- 30 incorrect predictions of death (False Positives)
- 18 incorrect predictions of survival (False Negatives)
- 56 correct predictions of death (True Positives)

To further evaluate the model's performance, additional metrics were used including accuracy, precision, recall and F1 score. These metrics assess the model's ability to accurately predict outcomes. Specifically, accuracy measures the overall prediction accuracy, precision measures the correct positive predictions, recall measures the correct identification of actual positive classes, and the F1 score combines recall and precision to provide a comprehensive evaluation. These metrics are calculated and presented in the below, providing a detailed assessment of the model's performance:

Table 8: Calculations of confusion matrix

METRICS	RESULTS
Accuracy	0.6000
Precision	0.6512
Recall	0.7568
F1 Score	0.7000

The above table shows that the Multi-Layer Perceptron (MLP) model achieved an accuracy of 60% in predicting maternal mortality from the test set. The model's precision was 65.12%, indicating that most of its positive predictions were correct. Additionally, the recall value of 75.68% shows that the model was able to identify most of the actual positive classes. Overall, the high F1 score suggests that the model performing well.

Ultimately, the model successfully predicted and classified maternal deaths, accurately identifying and categorizing the data. The results were categorized into true positives, true negatives, false positives (Type I errors), and false negatives (Type II errors). Table 8 provides a sample of 10 random observations from the original dataset, demonstrating the model's ability to predict maternal mortality from existing data.

Table 9: Predictions

Mortality	Predicted Mortality
1	0.7255562
3	0.5382390
7	0.5684380
12	0.4560086
15	0.4498486
18	0.4335789
22	0.4161610
27	0.4705353
28	0.4161610
32	0.6381355

Results were presented in the section analysed, and interpreted. The Multi-Layer Perceptron model was executed and examined on how best it fits the maternal data. The overarching aim of this study was to construct a predictive model that can predict and classify the misclassification of dead patients as alive, precision proved to be a better indicator of model performance in that it had the precision of 65.12%. The researchers ascertained that the more data we have, the more accuracy we derive from the algorithm. The following section will flesh-out the recommendations based on the research carried out by the researchers; it will also highlight the conclusions reached from this study.

CONCLUSION AND RECOMMENDATIONS

This study aimed to develop a predictive model for maternal mortality, a pressing concern globally with significant demographic implications. Childbirth is not only a natural process but also crucial for a country's workforce. To tackle this issue, relevant literature was reviewed to identify key contributing factors. The study employed machine learning algorithms, specifically a Multi-Layer Perceptron Artificial Neural Network, to predict and classify maternal deaths, and evaluated the model's performance to provide insights for addressing this critical concern. Since the maternal mortality rates are increasing at Chitungwiza Hospital, the researchers recommend the following policy interventions:

- That every nation try to follow the Global targets goals that is, MMR of above 140 per 100 000 live births should not be attained by any country, global maternal mortality ratio should be minimized to 70 per 100 000 lives;
- The CCH health executive should provide sufficient and quality medical care during delivery especially in Caesar wards;
- The government of Zimbabwe should intervene and provide enough family planning services at CCH and employing qualified doctors with the help of MoCCH;

- the application of machine learning techniques such as ANN in other health institutions where an Artificial intelligence model can provide answers to the most significant factors that cause maternal deaths in a shortest possible time.

From research conducted, it has been ascertained that age and parity are variables that are causing high maternal mortality but from the researcher's results mode of delivery, cause of deaths, booked and marital status are the top four causes of maternal deaths. This made the researchers to recommend the health executive at CCH should improve the quality of obstetric care being provided at the hospital.

The researchers also contemplate a further study aimed examine the relationship between statistics and machine learning. Additionally, it would be beneficial to include more factors related to maternal mortality in the dataset during the study. These factors have an impact on all patients, but the way they affect different patients can vary, and taking their variability into account over time can alter the results when the analysis is conducted at different times of the day.

REFERENCES

Alemayehu, Y. K. *et al.* (2019). Quality of Intrapartum Care for Pregnant Women and Newborns in Health facilities in Ethiopia, Kenya, and Tanzania: A Systematic Review. *PLoS ONE, 14*(8), 1-17.

Azhar, M., Badshah, A., Kabir, N. and Batool, Z. (2018). An Empirical Study of Factors Influencing Maternal Mortality Rate in Developing Countries. *American Journal of Social Science Research*, 4(1), 1-8.

Box, G. E. P. *et al.* (2015). *Time Series Analysis: Forecasting and Contro*l. John Wiley & Sons, Hoboken, New Jersey.

Bryman, A. (2016). *Social Research Methods* (5th ed.). US, New York: Oxford University Press, Oxford.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. United States, MIT Press, Cambridge, MA.

Hair, J. F., Black, W. C., Babin, B. J. and Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Boston, MA, Cengage Learning.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York: Springer. Index-Annual-Report.pdf. Accessed April 13, 2023

Kibret, K. T., Chojenta, C. and Gresham, E. (2015). Maternal Health Interventions in Ethiopia: A Systematic Review. *African Journal of Reproductive Health*, *19*(3), 45- 56.

Kumar, R. (2019). *Research Methodology: A Step-by-step Guide for Beginners* (5th ed.). London, Sage Publications.

Little, R. J. A. and Rubin, D. B. (2019). *Statistical Analysis with Missing Data* (3rd ed.). Hoboken, NJ, John Wiley & Sons.

Makworo, T. T. (2019). Forecasting Maternal Mortality Rates at Chitungwiza Central Hospital in Zimbabwe Using Artificial Neural Networks. *BMC Research Notes, 12*(1), 1-7.

Munjanja, S. P. and Lindmark, G. (2013). Maternal Mortality in Zimbabwe: Changes and the Way Forward. *African Journal of Midwifery and Women's Health*, 7(1), 18-23.

Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review, 65*(6), 386-408.

Tebbs, J. M. (2018). *Introduction to Time Series Analysis*. CRC Press, New York, Boca Raton, FL.

Tura, G. and Fantahun, M. (2019). Modelling Relationship between Maternal Mortality and other Independent Variables in Many Countries. *PLoS ONE, 14*(10), e0222836.

United Nations Population Fund. (2008). *Adding It Up: The Costs and Benefits of Investing in Sexual and Reproductive Health 2008*. New York: UNFPA.

World Health Organization. (2015). Trends in Maternal Mortality: 1990 to 2015: Estimates by WHO, UNICEF, UNFPA, World Bank Group and the United Nations Population Division. Geneva: World Health Organization.

World Health Organization. (2019). Maternal Mortality Ratio (per 100,000 live births). Available online: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/maternal-mortality-ratio-(per-100-000-live-births). Accessed on 15 April 2023.

World Health Organization. (2019). Maternal Mortality. Available online: https://www.who.int/news-room/fact-sheets/detail/maternal-mortality. Accessed on 15 April 15.