International Journal of Recent Advances in Multidisciplinary Topics
Volume 5, Issue 8, August 2024
https://www.ijramt.com | ISSN (Online): 2582-7839

Predicting Pregnancy Delivery Outcomes using Machine Learning Algorithms

Gabriel Mbatha Ngao[1*], Abraham Mutua Matheka[2]
[1,2]*Department of Computing and Information Science, School of Pure and Applied Science, Kenyatta University, Nairobi, Kenya*

Abstract*: The World Health Organization (WHO) has recognized the unpredictability of pregnancy delivery outcomes, including stillbirth, miscarriage, induced abortion, and live births, as a significant issue. Accurate prediction of these outcomes is crucial for improving healthcare treatments for mothers and children. This research uses a dataset from the Kenya Demographic and Health Survey (KDHS) to explore machine learning algorithms for predicting these outcomes. Pregnancy delivery outcomes are complex and rely on various demographic factors, including the number of antenatal care visits, the respondent's age, wealth index, marital status, place of residence, geographic location, religious affiliation, and the overall count of births. The research aims to develop reliable machine learning models that accurately predict these outcomes by examining the complex interactions between demographic factors. Various approaches, including neural network algorithms (ANNs), support vector machines (SVMs), and generalized linear models (GLMs), will be employed to find the model that best fits the needs of understanding and predicting the range of pregnancy delivery outcomes while accounting for these factors. The study also aims to reveal important correlations between Kenyan population demographic characteristics and delivery outcomes, with preliminary studies suggesting that the Random Forest Gini score may be a key factor in determining delivery results.

***Keywords*: WHO, KDHS, ANC, Machine Learning, ANNs, SVMs, GLMs, Random Forest Gini Score.**

1. Introduction

The experiences of pregnancy and birth are transformative, affecting the mother and the child for a lifetime. The World Health Organization (WHO) emphasizes the importance of precise prediction of pregnancy delivery outcomes, which can be customized by medical professionals to enhance care for expectant mothers and promote successful births (Osborn et al., 2022). Variables such as pregnancy-related visits, geography, religious affiliation, age, marital status, educational achievement, socioeconomic position, and previous birth history greatly influence these outcomes. ANC checkups are essential for early identification of issues and warning signals during pregnancy and childbirth, allowing for timely therapies and improved outcomes for both mothers and children. Prenatal vaccination contributes to healthy pregnancy and delivery outcomes by protecting mothers and babies against diseases that can be avoided (Belachew et al., 2022). However, only 64% of expectant women worldwide obtain the required minimum of four ANC consultations. The WHO has created extensive recommendations to address this gap and regularly revises guidelines about pregnancy-related issues (World Health Statistics 2022, 2022).

A. *Problem Statement*

Predicting the outcome of a prenatal birth is a challenging task, despite advancements in healthcare. Current models often focus too much on individual factors, neglecting the intricate relationships between them. This makes it difficult for medical practitioners to accurately identify high-risk pregnancies, administer targeted therapies, and allocate resources efficiently (Middleton et al., 2020). The project aims to develop machine learning algorithms using a large dataset like the Kenya Demographic and Health Survey (KDHS) to provide accurate prediction models. These models will investigate the complex relationships between demographic characteristics, improving the ability to forecast prenatal outcomes and aid in identifying high-risk pregnancies and resource allocation (Kowsher et al., 2021).

B. *Objectives*

This study aims to identify demographic characteristics of pregnant women and develop a Machine Learning model to predict delivery outcomes based on these characteristics.

1. To investigate Machine Learning algorithm that are being used in predicting pregnancy delivery outcome by pregnant women using KDHS dataset.
2. To examine and determine factors associated with pregnancy delivery outcome among pregnant women using KDHS dataset.
3. To develop and evaluate a robust Machine Learning model capable of predicting pregnancy delivery outcome given demographic characteristics using KDHS dataset.

2. Literature Review

Historically, conception was a mystery, with mothers believed to be the source of fertility and children considered property of men. The Genesis account inspired extreme ascetic practices, which significantly impacted Western conceptions of birthing (Payne et al., 2022) . Post-World War II, home births were common in Western countries, but World War II led to a

*Corresponding author: gabriel.mbatha@gmail.com

spike in hospital deliveries and a shift in medical practitioners' focus on birthing risks. Cultural anthropological research shows that different cultures have different results from delivery, with customs influencing postpartum customs and classifying people according to position and gender (Gómez-Roig et al., 2021).

Factors affecting pregnancy delivery outcome include antenatal care visits (ANC), religious affiliation, education level, geographic location, and respondent age. Antenatal care is crucial for teenage girls and pregnant mothers, providing care for pregnancy-related disorders, risk identification, preventative interventions, and health education (Belachew et al., 2022). The Focused Antenatal Care (FANC) paradigm, recommended by the World Health Organization, reduces maternal death rates by raising awareness of potentially fatal illnesses. However, global initiatives face challenges due to societal, economic, and cultural restrictions (Osborn et al., 2022).

Education significantly impacts pregnancy outcomes, with higher levels of education leading to better delivery results and a lower likelihood of early birth (Trost et al., 2017). Geographic location significantly impacts unfavourable pregnancy outcomes, with women aged 20-34 more likely to give birth normally, while African American women have a lower likelihood (Holcomb et al., 2021). Adolescent pregnancy is a major risk factor for health issues like postpartum haemorrhage, eclampsia, anaemia, and uterine infections (Ramadhani Nainggolan et al., 2022).

A. Machine Learning Algorithm Used

This study aims to develop a machine learning classifier that can predict pregnancy delivery results during the prenatal period, helping decide whether to proceed with live birth, stillbirth, or miscarriage based on patient demographic data (Raja et al., 2021). Machine learning is particularly suitable for obstetrical research due to the abundance of patient data from electronic medical records. However, causality inference from treatment remains a challenge due to non-accessible and noise-free data (Rahmayanti et al., 2022).

- Logistic regression is a statistical technique that estimates the influence of factors on the probability of certain outcomes (Inyang et al., 2020).
- Support Vector Machines (SVMs) are flexible and useful for predicting pregnancy outcomes due to their low misclassification error rate (Chelsea & Prima Rosa, 2024).
- Random forests are high-prediction classifiers that prevent overfitting in decision trees, increase prediction by producing multiple trees, estimate testing errors, and detect class imbalances in data sets (Lindblad Wollmann et al., 2021). They are useful for continuous prediction, classification, high-dimensional information handling, producing reliable signals, accommodating missing values, and being tolerant of outliers (Fazzari et al., 2022).
- Artificial neural networks are computational methods used to predict problems, especially in medical contexts, consisting of connected nodes that mimic functional relationships (Chelsea & Prima Rosa, 2024). They learn from information patterns and improve prediction accuracy using a training set of known values.

3. Conceptual Framework

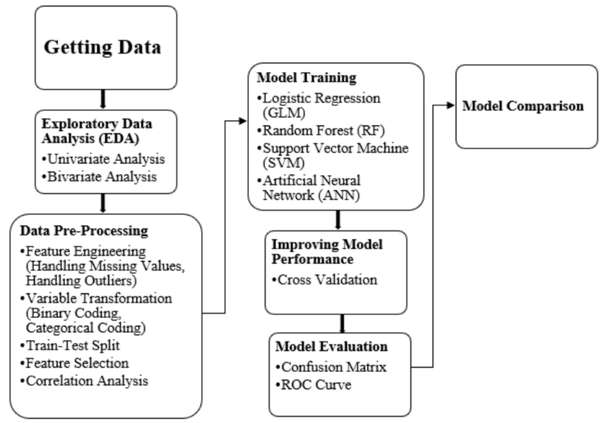

Fig. 1. Conceptual framework

4. Methodology

A. Overview

This section provides a detailed breakdown of the methodology, steps, and strategies used in this study to achieve its objectives.

B. Data Source

The KDHS conducted a national survey on mothers five years before the 2022 program, aiming to understand health indicators, knowledge, behaviour, and social concerns, including HIV, domestic abuse, and female genital mutilation.

C. Data Assessment for Exploration (EDA)

The research uses Exploratory Data Assessment (EDA) to analyse women's pregnancy delivery outcomes and demographic traits. It uses univariate and bivariate analysis techniques, including scatterplots and correlation matrices, to examine factors influencing outcomes, including prenatal care visits, location, and education (Hullman & Gelman, 2021).

D. Data Preliminary Processing Strategies

1) Feature Engineering

This research explores feature engineering techniques to enhance input data and predict accuracy in machine learning models. It aims to reduce overfitting, provide interpretive flexibility, and increase generalizability, ensuring robust models (X. Zhang et al., 2021).

a) Handling missing values

Missing data in data collection can lead to incorrect conclusions in machine learning tasks. Strategies like imputation, prediction modelling, and deletion are employed to minimize the impact of missing data. Imputation is used in this research to ensure dataset integrity and model stability, allowing for more reliable insights (Ma et al., 2021).

b) Handling outliers

Outlier detection is crucial in data mining and machine learning, identifying abnormal observations due to experimental mistakes or measurement flaws. Methods include Grubbs, Hampel filter, Dixon tests, and basic descriptive statistics (Alghushairy et al., 2020).

$$[q0.25 - 1.5IQR, q0.75 + 1.5IQR]$$

2) Variable transformation

a) Binary Coding

To improve prediction models for pregnant delivery outcomes, the researcher used variable transformation techniques. They categorized pregnant women based on factors such as age, wealth index, marital status, education level, and birth number, using binary coding to simplify interpretation (H. Zhang & Wu, 2023).

$$X_{(k)} = \begin{cases} 1, \textit{ if } X \geq C_k; \\ 0, \textit{ if } X < C_k; \end{cases} \tag{1}$$

The researcher aimed to convert the target variable into a binary format for a comprehensive study on the complex interactions between detected parameters and pregnancy delivery probability.

b) Categorical Encoding

Machine learning models enhance accuracy by converting textual input into numerical values using techniques like binary encoding, one-hot encoding, and label encoding. One-hot encoding produces a final vector space with evenly distanced and orthogonal categories, while base levels are chosen, and dummy variables are added for each category (Catillo et al., 2022).

Table 1
One: Formulation of hot encoding

Z	Z_B	Z_C	Z_D
A	0	0	0
B	1	0	0
C	0	1	0
D	0	0	1

3) Train-Test Slit

The train-test divide is a technique used to evaluate a model's efficacy by dividing research data into training and test sets. The training set helps identify correlations and patterns, while the test set measures generalization. A 70%-30% train-test split threshold is used to assess model performance across various machine learning applications (Thapa et al., 2022).

4) Random Forest Based Feature Selection and Significance

Feature selection is crucial in data science operations, particularly for large datasets. Random Forests are popular due to their tree-based methods, which improve node purity, simplify models, reduce variance, prevent overfitting, and reduce training time and computational expenses (Lindblad Wollmann et al., 2021). The variable importance approach ranks variables based on prediction performance, aiming to maximize prediction performance and improve processing efficiency. The Gini impurity, or i(τ), is a crucial measure in determining the best split at each node in a random forest's binary trees. It measures the effectiveness of a possible split in separating samples from two classes within a node, serving as an effective gauge of entropy (Fazzari et al., 2022). The calculation considers the percentage $P_K = \frac{n_K}{n}$ of samples belonging to class k = {0, 1} among all n samples at node τ

$$i(\tau) = 1 - P_1^2 - P_0^2 \tag{2}$$

The decrease, represented by Δi, that results from dividing samples and allocating them to τ_i and τ_r sub nodes by means of a threshold t_θ applied to variable θ is shown as a function of the corresponding sample proportions ($P_i = \frac{n_i}{n}$ and $P_r = \frac{n_r}{n}$).This formulation encapsulates the quantitative effect of the separation process and sheds light on how the distribution of samples among the generated sub nodes is affected by the threshold that is selected for the given variable.

$$\Delta i(\tau) = i(\tau) - P_i i(\tau_i) - P_r i(\tau_r) \tag{3}$$

The ideal split reduces Gini impurity, $\Delta i\theta$ (τ, T), by recording and averaging it over all nodes in each tree T in the Random Forest ensemble. This process assesses the efficacy of different splits and their effectiveness in improving node purity and predictive performance.

$$I_G(\theta) = \sum_T \sum_\tau \Delta i\theta\,(\tau, T) \tag{4}$$

This score shows the degree of general discriminative usefulness and the frequency with which a certain feature was chosen for a split for the given classification task.

5) Correlation

The correlation technique, Pearson correlation approach, was utilized to assess the degree and orientation of association between two variables, focusing on linear dependency (Bharat et al., 2020).

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum((x_i - \bar{x})^2 \sum(y_i - \bar{y})^2} \tag{5}$$

The value of the collection's first and second variables are denoted by x_i and y_i , respectively, and *r* is the Pearson product moment correlation coefficient. The dataset's first and second variable means are, respectively, $\bar{x}$ and ȳ.

E. Machine Learning Models Used

1) Logistic regression

A vital statistical analysis method is logistic regression, which calculates the probability of a certain occurrence using a generalized linear model (GLM). It will be widely used due to its effectiveness and ease of use, and its quick data scoring speed and high computational efficiency make it a popular choice in various analytical contexts (Panda et al., 2022).

$$P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_q X_q \tag{6}$$

The probability (P) it will be determined by integrating predictor variables into a logistic response function, ensuring it remains within 0 and 1. The odds ratio, the ratio of success to non-success, is also considered, creating an exponential equation.

$$\textit{Odds ratio}\ (\mathrm{Y} = 1) = \frac{P}{1-P} \tag{7}$$

$$\mathrm{P} = \frac{Odds}{1+\ Odds} \tag{8}$$

With the logistic response function combined, we produce.

$$Odds(Y = 1) = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_q X_q} \tag{9}$$

Taking Logarithm,

$$\log(Odds(Y = 1) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_q X_q \tag{10}$$

2) Random Forest

Random Forest (RF), a supervised tree-based classification model, effectively solves classification and regression issues using decision trees, even when predictive features contain irrelevant variables (Sevgen, 2020). Its versatility and adaptability are emphasized by its unique bootstrap aggregation method.

Training set is given X = $X_1, X_2, \ldots, X_n$.
Response variable Y = $Y_1, Y_2, \ldots, Y_n$

The process of bagging entails training trees on a random sample that will be chosen with replacement from the training dataset repeatedly (*B* times) (C. Zhang et al., 2021).

Using replacement, a sample of n training instances from datasets *X* and *Y* is taken for every iteration $b = 1, 2 \ldots B$. The names X_b and Y_b, respectively, refer to these sampled datasets.

A regression or classification tree is trained by the model. f_b on X_b, Y_b

To acquire predictions for unknown samples x', the process will aggregate the predictions from each individual regression tree on x'; for classification trees, it will take the majority vote after training.

$$\dot{f} = \frac{1}{B} \sum_{b=1}^{B} f_b(x') \tag{11}$$

By minimizing model variance without raising bias, this bootstrapping technique improves the performance of models.

3) Support Vector Machine (SVM) algorithms

Support Vector Machine (SVM) employs a hyperplane for homogeneous data division, forecasting intricate relationships using regression modelling and K-Nearest Neighbour. It efficiently divides linearly separable training data, positioning the margin hyperplane for best separation.

$$\vec{\mathrm{W}}.\vec{\mathrm{x}} - b = 1\ and \tag{12}$$
$$\vec{\mathrm{W}}.\vec{\mathrm{x}} - b = -1 \tag{13}$$

These two hyper planes should ideally be equally spaced apart.

$\frac{2}{||\mathrm{W}||}$ To extend the separation between the planes, we want to reduce ||W ||. To prevent data points from falling outside of the margin, we may apply the following restriction to each i.

$$\vec{\mathrm{W}}.\vec{\mathrm{x}} - b \geq 1\ if\ \ \mathrm{y_i} = 1\ or \tag{14}$$

$$\vec{\mathrm{W}}.\vec{\mathrm{x}} - b \geq -1\ if\ \ \mathrm{y_i} = -1 \tag{15}$$

These rules state that each data point must be on the proper side of the margin.

This may be expressed mathematically as:

$$y_i(\vec{\mathrm{W}}.\vec{\mathrm{x}} - b) \geq 1\ forall\ 1 \leq i \leq n \tag{16}$$

When we combine them, we get the following efficiency problem:

Reduce the value of ||W|| subject to $y_i(\vec{\mathrm{W}}.\vec{\mathrm{x}} - b) \geq 1\ for\ i = 1, \ldots, n$

Where $\vec{\mathrm{W}}$ and b that determines the classifier and resolves this issue $\vec{x} \mapsto sign(\vec{\mathrm{W}}.\vec{\mathrm{x}} - b)$

4) Artificial Neural Network algorithm (ANN)

To find patterns and connections in data, artificial neural networks (ANNs) use biology and the functioning of the human brain. Multilayer networks—also referred to as perceptron's—identify complicated associations by using input nodes to convey hidden layers. The hidden-layer computation process introduces a bias constant for improved model accuracy, while the activation function introduces nonlinearity for complex interactions. Forward propagation and backpropagation rectify significant faults (Gómez-Roig et al., 2021). The thresholds for parameters and α, together with the weights $w_1, w_2, \ldots, w_n$, are all randomly assigned integer values. For the perceptron to produce the required output, $Y_d(p)$, it must be fed the inputs $x_1(p)$, $x_2\ (p) \ldots\ x_n\ (p)$, where iteration p denotes the perceptron's P^{th} training example (p = 1, 2...). In the procedure, the real output *Y* (p) is calculated at iteration p = 1.

$$Y(p) = step[\sum_{k=0}^{n} X_i(p) W_i(p) - \theta\ ] \tag{17}$$

i.e.

$$\mathrm{Y(X)} = \mathrm{step}[\mathrm{X}] = \begin{cases} 1,\ if\ x \geq 0 \\ 0,\ if\ x < 0 \end{cases}$$

In this instance, *n* indicates the number of inputs for a backpropagation, while step denotes the step activation function.

Next, the feedforward's weights are adjusted.

$$W_i(p+1) = W_i(p) + \Delta W_i(p)$$

Where ΔW_i (p), the weight adjustment at iteration p, is determined using the delta rule.

The following provides the delta rule: $\Delta W_i(p) = \alpha$ x $X_i(p)(p)$

For each p between 1 and 3, we get e(p) = Yd(p) − Y (p), where α represents the learning rate and α⊆ (0, 1).

Every time reiterating p is increased by 1, convergence is attained, that is, error e = 0. eights or coefficients.

F. Increasing the Model Efficiency

1) Cross Examination

a) Cross examination in K-Fold

For performance assessment, a dataset is divided into k equal-sized folds using the K-fold Cross-Validation technique. The test set for each iteration consists of one-fold, while the training set consists of the remaining k-1 folds. With the data from the test set, the Mean Squared Error (MSE) is computed. Using a new validation set of data each time, the iterative process is repeated k times (Saud et al., 2020). The K-fold Cross-Examination estimate is computed using the average MSE values.

$$CV_k = \frac{1}{k}\sum_{k=1}^{k} MSE_i \quad (18)$$

For the categorization scenario

$$Error\ rate\ in\ training = \frac{1}{n}\sum_{i=1}^{n} I(y_i \neq \hat{y}_i) \quad (19)$$

In the context of K-fold Cross-Validation extension, $I\ (y_i \neq \hat{y}_i)$ is an indicator variable that equals 0 when $y_i = \hat{y}_i$ and 1 when $y_i \neq \hat{y}_i$. y_i represents the actual class label for the ith observation. The i^{th} observation is marked as successfully anticipated if $I\ (y_i \neq \hat{y}_i) = 0$.

G. Metrics for Analyzing Models

1) Confusion matrix

Four distinct prediction possibilities are included in the confusion matrix in Figure 3.8.1, which depicts a two-class classification scenario. False positives and false negatives indicate mistakes in categorization, while true positives and true negatives demonstrate accurate results. According to their position on the confusion matrix, the research defines terms and presents a thorough evaluation of the model's efficacy by displaying the ratio of correctly classified words to various kinds of incorrectly classified words (Xu et al., 2020).

Table 2
Illustration of the Confusion Matrix

		Estimated class	
		Inverted	Positive
Real-time Class	**Inverted**	*A*	*B*
	Positive	*C*	*D*

a) True Negative (TN), the correct predictions when the actual result is negative.
b) False Positive (FP), the inaccurate forecasts in which instances are categorized as positive.
c) False Negative (FN) denotes incorrect forecasts in situations considered as negative.
d) True Positive (TP) is the accurate forecasts when the result turns out to be positive.

Using a formula, the percentage of true positives and negatives compared to the total number of cases determines a system's accuracy.

$$Accuracy = \frac{a+b}{a+b+c+d} = \frac{TN+FP}{TN+FP+FN+TP}$$

The true positive rate is defined as the percentage of accurately classified positive cases.

$$Recall\ or\ sensitivity\ (True\ Positive\ Rate) = \frac{d}{c+d} = \frac{TP}{TN+FP}$$

The percentage of instances that were incorrectly classified as positive when they were negative is known as the false positive rate.

$$False\ Positive\ Rate = \frac{b}{a+b} = \frac{FP}{TN+FP}$$

The percentage of correctly identified negative instances was used to establish the genuine negative rate:

$$True\ Negative\ Rate = \frac{a}{a+b} = \frac{TN}{TN+FP}$$

The proportion of positive instances that were mistakenly classified as negative is known as the false negative rate.

$$False\ Negative\ Rate = \frac{c}{c+d} = \frac{FN}{FN+TP}$$

Finally, the fraction of correct forecasted positive instances is displayed by precision, sometimes referred to as positive predictive value:

$$Precision = \frac{d}{b+d} = \frac{TP}{FP+TP} \quad (20)$$

2) Receiver Operating Characteristics Curve (ROC)

The ROC graph is a visual diagnostic tool that evaluates a classifier's effectiveness in categorizing positive and negative examples. It displays the true positive rate (Y-axis) and the false positive rate (X-axis). The ROC curve provides insights into the classifier's discriminative skills, allowing for a nuanced understanding of the trade-off between sensitivity and specificity across different classification thresholds (Luckett et al., 2021). This method helps choose an optimal operating point based on goals and limitations, improving the interpretability of classifier performance.

5. Conclusion

This study developed a machine learning model to predict pregnancy delivery outcomes based on demographic features from the KDHS dataset. The model used a Random Forest classifier to rank demographic attributes related to ANC visits, including location, religious affiliation, wealth index, education level, and total number of children born. The model was then fitted to a binary logistic regression model to determine the relative contributions of each feature. The research found that ANC visits, wealth index, education level, a woman's current age, supportive spouse, number of children born to date, and delivery location are important factors affecting pregnancy delivery outcomes. The algorithm alert expectant mothers at risk of not receiving healthcare services to prevent complications and lower maternity-related mortality rates. Future studies could include putting the model online for medical practitioners to use.

References

[1] Alghushairy, O., Soule, T., Alsini, R., & Ma, X. (2020). A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams.
[2] Bharat, G. R., Subhashini, M., & Gopinath, R. (2020). Employee Attrition Prediction in Industry Using Machine Learning Techniques. International Journal of Advanced Research in Engineering and Technology, 11(12), 3329–3341.
[3] Chelsea, M. Y., & Prima Rosa, P. H. (2024). Classification of delivery type of pregnant women using support vector machine. E3S Web of Conferences, 475, 02015.
[4] Fazzari, M. J., Guerra, M. M., Salmon, J., & Kim, M. Y. (2022). Adverse pregnancy outcomes in women with systemic lupus erythematosus: can we improve predictions with machine learning? Lupus Science & Medicine, 9(1), e000769.
[5] Gómez-Roig, M. D., Pascal, R., Cahuana, M. J., Garcia-Algar, O., Sebastiani, G., Andreu-Fernández, V., Martínez, L., Rodríguez, G., Iglesia, I., Ortiz-Arrabal, O., Mesa, M. D., Cabero, M. J., Guerra, L., Llurba, E., Domínguez, C., Zanini, M. J., Foraster, M., Larqué, E., Cabañas, F., Vento, M. (2021). Environmental Exposure during Pregnancy: Influence on Prenatal Development and Early Life: A Comprehensive Review. Fetal Diagnosis and Therapy, 48(4), 245–257.
[6] Inyang, U., Inyang, U. G., Eyoh, I. J., Nwokoro, C. O., Osang, F. B., & Afolorunso, A. A. (2020). Comparative Analytics of Classifiers on Resampled Datasets for Pregnancy Outcome Prediction. Article in International Journal of Advanced Computer Science and Applications, 11(6).
[7] Kowsher, M., Jahan Prottasha, N., Tahabilder, A., Abdur Rakib, M., Abdur-Rakib, M., Shameem Alam, M., & Habib, K. (2021). Predicting the Appropriate Mode of Childbirth using Machine Learning Algorithm. Article in International Journal of Advanced Computer Science and Applications, 12(5), 2021.
[8] Lindblad Wollmann, C., Hart, K. D., Liu, C., Caughey, A. B., Stephansson, O., & Snowden, J. M. (2021). Predicting vaginal birth after previous cesarean: Using machine-learning models and a population-based cohort in Sweden. Acta Obstetricia et Gynecologica Scandinavica, 100(3), 513–520.
[9] Luckett, D. J., Laber, E. B., El-Kamary, S. S., Fan, C., Jhaveri, R., Perou, C. M., Shebl, F. M., & Kosorok, M. R. (2021). Receiver Operating Characteristic Curves and Confidence Bands for Support Vector Machines. Biometrics, 77(4), 1422–1430.
[10] Middleton, P. G., Gade, E. J., Aguilera, C., MacKillop, L., Button, B. M., Coleman, C., Johnson, B., Albrechtsen, C., Edenborough, F., Rigau, D., Gibson, P. G., & Backer, V. (2020). ERS/TSANZ Task Force Statement on the management of reproduction and pregnancy in women with airways diseases. European Respiratory Journal, 55(2).
[11] Rahmayanti, N., Pradani, H., Pahlawan, M., & Vinarti, R. (2022). Comparison of machine learning algorithms to classify fetal health using cardiotocogram data. Procedia Computer Science, 197, 162–171.
[12] Raja, R., Mukherjee, I., & Sarkar, B. K. (2021). A Machine Learning-Based Prediction Model for Preterm Birth in Rural India.
[13] Saud, S., Jamil, B., Upadhyay, Y., & Irshad, K. (2020). Performance improvement of empirical models for estimation of global solar radiation in India: A k-fold cross-validation approach. Sustainable Energy Technologies and Assessments, 40, 100768.
[14] Thapa, C., Arachchige, P. C. M., Camtepe, S., & Sun, L. (2022). SplitFed: When Federated Learning Meets Split Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 36(8), 8485–8493.
[15] T, Tewabe., A, Belachew., & G, Dessie. (2022). Neonatal mortality and its association with antenatal care visits among live births in Ethiopia: a systematic review and meta-analysis. The Journal of Maternal-Fetal & Neonatal Medicine.
[16] Villano, U., A, Pecchia., & Catillo, M. (2022). AutoLog: Anomaly detection by deep autoencoding of system logs. Expert Systems with Applications.
[17] Daniel, K., Peter, O.S., Emannuel, G., Martin, Obeagu, E. I., & Daniel, K. (2023). Factors Influencing Early Neonatal Adverse Outcomes among Women with HIV with Post Dated Pregnancies Delivering at Kampala International University Teaching Hospital, Uganda.
[18] Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. Journal of Experimental Psychology: General.
[19] Honkavuo, L. (2021). Women's experiences of cultural and traditional health beliefs about pregnancy and childbirth in Zambia: An ethnographic study. Health Care for Women International.
[20] Gelman, A., & Hullman, J. Designing for Interactive Exploratory Data Analysis Requires Theories of Graphical Inference.
[21] Fotouhi, A., M, Yunesian., Hussein, H., M. S, Hasanvand., Shamsipour, M., & Hasanvand, M. S. (2017). Association of adverse birth outcomes with exposure to fuel type use: A prospective cohort study in the northern region of Ghana.
[22] Althoffy, K. N., Leifheit, K. M.,Black, M.M., Schwartz, G. L., Edin, K.J., Pollack, C. E., & Jennings, J. M. (2020). Severe Housing Insecurity during Pregnancy: Association with Adverse Birth and Infant Outcomes.
[23] Peng, X., Ma, M.,Wu., C., Tulyakov, S., Ren, J., Zhao, L., &Tulyakov, S.(2021). SMIL: Multimodal Learning with Severely Missing Modality.
[24] Haller, T., Nyambedha, E., Okumu Ombere, S., & Merten, S. (2021). Cultural Practices During Pregnancy and Birth Among the Giriama Community in Coastal Kenya.
[25] Linscott, L. L., Poussaint, T. Y., Osborn, A. G., Louis, D. N., & Salzman, K. L. (2022). The 2021 World Health Organization Classification of Tumors of the Central Nervous System.
[26] N. R. Panda, J. K. Pati, J. N. Mohanty, & R. Bhuyan (2022). A Review on Logistic Regression in Medical Research.
[27] Brown, L.J., Payne, E., Rollo, M., Ryan, K., Harrison, S., Collins, C., Schumacher, T., & Wynne, K. (2022). Rural healthcare delivery and maternal and infant outcomes for diabetes in pregnancy:
[28] E1, L., Rahmi, L., Jurnal., & Bustami, S. (2021). The Relationship of The Characteristics of The Mother with The Type of Delivery.
[29] Ramadhani Nainggolan, Riska Yanti Harahap, & Ratna Wulandari. (2022). The Influence of Health Education About Delivery on Primigravidal Anxiety Trimester III During the Covid-19 Pandemic in Aek Haruaya Village.
[30] Trost, S., Harvey, A., Goodman, D.A, Beauregard, J., Chandra, G., Njie, F., Berry, J., & Harvey, A. (2017). Pregnancy-Related Deaths: Data from Maternal Mortality Review Committees.
[31] World Health Statistics 2022. (2022).
[32] Zhang, X.I.N., WU, X., Zhang, C., Wang, X., Chen, S., Li, H., & Zhang, X., (2021). A Modified Random Forest Based on Kappa Measure and Binary Artificial Bee Colony.
[33] Wang, Q., Xu, W., Han, Y., & Zhang, X. (2021). HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture.