FORECASTING VARIOUS CHILDBIRTH METHODS USING MACHINE LEARNING

Yazhini S
Assistant Professor of Information Technology
Nandha Engineering College-Erode
Yazhini@nandhaengg.org

Deepan M S
Information Technology
Nandha Engineering College-Erode
Deepan.20it008@nandhaengg.org

Karunagaran A
Information Technology
Nandha Engineering College-Erode
karunakaran.20it018@nandhaengg.org

Niranjan Kumar S V
Information Technology
Nandha Engineering College-Erode
niranjankumar.20it034@nandhaengg.org

ABSTRACT-Childbirth is a multifaceted process influenced by various factors, including maternal health, fetal conditions, and labor dynamics. Accurate prediction of the mode of delivery, whether cesarean section or vaginal birth, is pivotal for ensuring optimal maternal and neonatal outcomes. We employed a combination of Support Vector Machines (SVMs) and Random Forests (RFs), two powerful machine learning algorithms, to enhance prediction accuracy. SVMs excel in classifying data into multiple classes, while RFs, an ensemble learning algorithm, integrate predictions from multiple decision trees to produce more robust forecasts. Our results indicate that the integration of SVMs and RFs significantly improves the accuracy of predicting the mode of childbirth compared to individual algorithms. The combined approach of SVMs and RFs effectively captures the complexity of childbirth data, offering valuable insights for clinicians to make informed decisions. This predictive model not only enhances clinical decision-making but also empowers expectant mothers by providing insights into their birthing options. By advancing the capabilities of predictive analytics, this study contributes to improving healthcare quality, reducing unnecessary interventions, and enhancing maternal and neonatal safety.

Keywords: Childbirth Prediction, Mode of Delivery, Support Vector Machines (SVMs), Random Forests (RFs), Machine Learning Algorithms, Maternal Health, Neonatal Outcomes, Clinical Decision-Making, Predictive Analytics, Obstetrics, Healthcare Quality, Maternal Safety, Neonatal Safety, Informed Decision-Making, Ensemble Learning.

1. INTRODUCTION

The use of machine learning algorithms in the healthcare industry has grown significantly in recent years, providing novel solutions to challenging problems. Predicting the modes of childbirth is one such crucial area of study, where identifying important characteristics is essential. Researchers and practitioners are investigating different machine learning algorithms to identify the most pertinent aspects that lead to effective predictions in delivery outcomes as the healthcare industry continues to harness the power of data. The desire to improve overall healthcare outcomes, optimize resource allocation, and improve maternity and neonatal care is what motivates this investigation. Under these circumstances, the combination of cutting-edge computational methods and obstetric data presents enormous potential for identifying trends and connections that could help inform choices made in the delivery room. This research explores the field of machine learning with the goal of determining and assessing the most significant characteristics that lead to the precise prediction of delivery modes, therefore opening the door to more individualized and successful obstetric healthcare interventions.

1.1 CHILDBIRTH PREDICTION

Predicting the result of childbirth is a critical area in modern medicine, where the incorporation of cutting-edge technologies has started to reshape conventional methods. The use of machine learning and predictive modeling algorithms has become a game-changing method in obstetrics in recent years. In order to predict the delivery method, a complex analysis of various aspects affecting the health of the mother and the newborn must be done. Scholars and medical professionals want to decipher complex patterns in obstetric data by utilizing data analytics and computational methodologies, which will ultimately lead to better informed decisions being made in the delivery room. Predicting childbirth modes with sufficient accuracy not only presents the potential to optimize resource allocation and improve healthcare outcomes, but it also highlights the possibility of customized and individualized interventions aimed at making every childbirth experience as safe and optimal as possible.

1.2 MODE OF CHILDBIRTH

One of the most important aspects of maternity care is the mode of delivery, which symbolizes the various methods that a woman might give birth. These complex phenomena include a range of delivery techniques, such as cesarean procedures and vaginal births, each with unique effects on the health of the mother and the newborn. Healthcare professionals and expectant women must be able to predict and understand the mode of birthing since it affects how prenatal care, labor, and postpartum recovery proceed. The best mode is determined by a complex interplay of factors including fetal appearance, mother health, and obstetric history. It is critical to investigate prediction models and cutting-edge technology in order to guarantee the health of mother and child. The present study aims to investigate the various aspects related to delivery technique, investigating the complex interactions between medical, societal, and personal elements that influence this essential part of the reproduction process.

1.3 MACHINE LEARNING

Within the field of artificial intelligence (AI), machine learning focuses on creating models and algorithms that allow computers to learn from data and make judgments or predictions without explicit programming. Put differently, machine learning makes it possible for machines to learn from experience and get better at it, just like people do. supervised learning, unsupervised learning, and reinforcement learning are some of the several methods and algorithms used in machine learning. A computer algorithm is trained via supervised learning using a labeled dataset, in which every data point has a corresponding output or response labeled on it. When an algorithm is trained on an unlabeled dataset, it must independently find patterns and relationships in the data.

1.4 HEALTHCARE

The dynamic and essential component of contemporary society that is healthcare includes a wide range of services aimed at maintaining and enhancing the well-being of the individual and the community. Healthcare is an intricate web of interconnected medical knowledge, cutting-edge technology, and humane care, based on the concepts of prevention, diagnosis, treatment, and rehabilitation. It penetrates research projects, public health campaigns, and the never-ending search for new ways to tackle pressing health issues. It is not limited to hospitals and clinics. The importance of healthcare is not limited to its ability to treat illnesses; it also serves as a foundation for creating a community in which people can live happy, healthy lives. The healthcare environment is constantly changing in this era of tremendous developments in medical science and rising emphasis on patient-centered approaches. This introduction lays the groundwork for a discussion of the complex and dynamic nature of healthcare, highlighting its vital role in determining the general well-being of people and communities.

1.5 MATERNAL AND NEONATAL CARE

In the field of medicine, maternal and neonatal care is fundamental, protecting the health and welfare of women and their infants during the crucial stages of pregnancy, childbirth, and the postpartum period. Recognizing the significant influence that mother health issues have on the neonate, this specialized profession is propelled by a holistic approach that gives priority to the physical, mental, and social elements of maternal health. In order to ensure a safe and healthy transition for mother and child, maternal and neonatal care covers a maximum of services, from prenatal care and labor and delivery aid to postpartum support. Continuous research and technological advancements drive innovation, improving outcomes and enhancing the quality of life for mothers and their newborns.

2. LITERATURE SURVEY

In this paper, Nafiz Imtiaz Khan [1] et al. make the proposal that the rising incidence of cesarean sections worldwide has sparked worries because of the possible short- and long-term consequences that may arise for women. Although in certain cases of problems after childbirth, this surgical treatment can save lives, overuse of it has hazards. The goal of this project is to use data mining techniques to anticipate when a cesarean section will be necessary in order to solve this issue and ultimately improve the safety of women and newborns by preventing needless surgery. Three ensemble prediction models based on the XGBoost, AdaBoost, and Cat boost algorithms were used in the study. According to the findings, XGBoost had the best accuracy at 88.91%, closely followed by AdaBoost at 88.69% and Cat boost at 87.66%. These models present a promising way to forecast the necessity of cesarean sections, enabling more informed choices to be made throughout labor and delivery. Interestingly, the study pinpointed several important characteristics that greatly enhance the prediction models' accuracy. The variables that were shown to be most important in properly predicting the chance of a cesarean section were amniotic fluid levels, medical indications, fetal intrapartum pH, the number of prior cesareans, and pre-induction factors. By emphasizing these significant characteristics, the study offers insightful information that might help medical professionals decide on delivery procedures with greater knowledge.

In a study conducted by Alberto De Ramón Fernández, Daniel Ruiz Fernández, and María Teresa Prieto Sánchez, the potential of three popular artificial intelligence algorithms—Support Vector Machines, Multilayer Perceptron, and Random Forest—was explored to develop a clinical decision support system for predicting the mode of delivery. The study utilized a comprehensive clinical database comprising 25,038 records with 48 attributes of women who gave birth at the Service of Obstetrics and Gynecology of the University Clinical Hospital "Virgen de la Arrixaca" in the Murcia Region (Spain) between January 2016 and January 2019. The participants were patients with singleton pregnancies who presented to the emergency room in active labor or underwent planned induction of labor for medical reasons. The findings revealed that all three algorithms demonstrated similar performance, achieving an accuracy rate of 90% or higher in classifying between caesarean and vaginal deliveries. However, the accuracy was slightly lower, around 87%, in distinguishing between instrumental and euthocic deliveries. These results validate the potential of using artificial intelligence algorithms as effective tools for developing clinical decision support systems to aid gynecologists in predicting the mode of delivery accurately.

Ping Guan et al. conducted a study to analyze the maternal and fetal factors influencing EmCS and develop a risk scoring system for its quantitative prediction. The study included 10,295 pregnant women, with 9.63% undergoing EmCS due to failed vaginal delivery, predominantly caused by fetal distress (67.21%) and abnormal fetal position (14.93%). Significant differences were observed in 17 maternal and fetal factors between the normal vaginal delivery (NVD) and EmCS groups. Based on these factors, a risk scoring system with six grades was established, demonstrating high discriminative power in predicting EmCS risk. The rate of EmCS increased progressively across the risk grades, ranging from 1.30% in grade 1 to 39.71% in grade 6. The system exhibited good predictive ability with an area under the curve of 0.787.

According to ASWAN SUPRIYADI SUNGE [3] et al.'s proposal in this paper, giving birth is a life-changing experience for every woman following pregnancy, with a range of dangers from difficult delivery to the mother's possible death. Using Caesarean sections strategically is one way to reduce these risks. In this situation, ML applications become an important tool for determining whether a Caesarean section is required for expectant moms. The dataset that was used for this study is composed of 3602 instances of 32 attributes that were obtained from outside data sources. To determine whether a Caesarean section is necessary, the study uses models from Random Forest, Decision Tree, and XGBoost. The goal of the comparison study is to determine which model, when all aspects are taken into account, performs best overall. The study also investigates how well the models function when particular attributes are chosen for examination. The results show that when all features are taken into account, Random Forest gets the maximum accuracy (0.86). But XGBoost performs best (0.85) when specific variables are taken into account, like age, obesity, and the number of days until the next pregnancy.

MUHAMMAD NAZRUL ISLAM et al. conducted a study to identify key features and employ machine learning algorithms for accurate prediction. Analyzing 6157 birth records, they identified 32 relevant features for childbirth prediction, categorizing them based on importance. Among the five machine learning algorithms tested, stacking classification (SC) achieved the highest f1 score of 97.9%, followed by random forest (RF) at 97.3%. The k-nearest neighbors (KNN) method scored 95.8%, while decision tree (DT) and support vector machine (SVM) achieved 93.2% and 88.6% respectively. These results highlight the potential of machine learning in enhancing delivery mode prediction accuracy, for optimal maternal and neonatal outcomes, create a decision that is far more trustworthy.

In this research, Kowsher [5] et al. suggest that choosing the optimal baby delivery techniques is essential to safeguarding the newborn child and the mother. However, while making this choice in a computerized manner, it is still necessary to investigate the optimum combinations of attributes. For this reason, we attempt to use AI to identify the most effective way to deliver babies. These days, computerized computation models such as deep learning and machine learning are widely employed in medical decision-making. Here, we have employed binary classification algorithms based on machine learning to help choose between two delivery techniques.

The prediction of vaginal birth after cesarean (VBAC) deliveries remains a complex challenge in obstetric care, influencing patient-physician decision-making and the rates of trials of labor. Ms. Michal Lipschuetz et al. aimed to develop a personalized prediction tool for VBAC using machine learning algorithms to aid in clinical decision-making. Their study yielded a personalized prediction tool with an area under the curve (AUC) of 0.745 (95% CI=0.728-0.762).

Data available at the first antenatal visit (first trimester model). This AUC improved to 0.793 (95% CI=0.778-0.808) with the inclusion of delivery unit admission features (pre-labor model). Using the pre-labor model, the majority of parturients (67.9%) were classified into low- and medium-risk groups with high VBAC success rates of 94.9%. In contrast, the high-risk group, comprising 32.1% of parturients, had a lower VBAC rate of 32.1%. Notably, the VBAC success rates remained high at 88% even in a population with a high rate of trial of labor after cesarean (TOLAC) attempts (75.6%). This study serves as a proof of concept for the application of machine learning algorithms in clinical decision-making regarding VBACs, demonstrating the potential of personalized prediction tools to improve VBAC success rates and guide patient-specific care strategies.

3. EXISTING SYSTEM

The rising global cesarean section (CS) rate presents both benefits and risks to maternal and neonatal health. Current CS decision-making often relies on clinical judgment, leading to potential unnecessary surgeries and associated risks. Data mining and machine learning techniques have been explored to improve CS prediction accuracy. Our project aims to develop ensemble prediction models using XGBoost, AdaBoost, and CatBoost algorithms, along with traditional Decision Tree and KNN algorithms. Key factors influencing CS prediction include amniotic fluid levels, medical indications, fetal intrapartum conditions, prior cesarean history, and pre-induction processes. The project seeks to enhance prediction accuracy to reduce unnecessary CS and improve maternal and neonatal safety during childbirth.

4. PROPOSED SYSTEM

One of the most important factors in guaranteeing the safety of mother and child is the manner of delivery. The attending physician's view is currently used to anticipate the mode of birth; nevertheless, selecting the incorrect delivery style might result in various short- and long-term health complications for both mother and child. The suggested method for predicting the mode of delivery through the use of random forests (RFs) and support vector machines (SVMs) is a novel strategy that could enhance the outcomes, lower expenses, and increase accuracy of childbirth. Pregnant women's data is gathered, preprocessed, important characteristics are chosen, SVM and RF models are trained, assessed, and the models are put into production. The system is relatively simple to deploy and use, and it is anticipated to attain higher accuracy and greater generalizability than conventional methods of forecasting the mode of birthing.

4.1 LOAD DATA

Data on delivery cases has been loaded into the dataset, capturing various essential characteristics like medical history, gestational age, maternal age, and prenatal care. This extensive dataset serves as a comprehensive repository of information crucial for our analysis. To ensure the accuracy and reliability of our findings, we meticulously reviewed the data for any anomalies or inconsistencies. Any missing values or errors were addressed and corrected to maintain the integrity of the dataset. This rigorous data cleaning process is vital as it lays the groundwork for accurate and meaningful analysis. Ensuring data quality is paramount, as it directly impacts the validity and trustworthiness of our results.

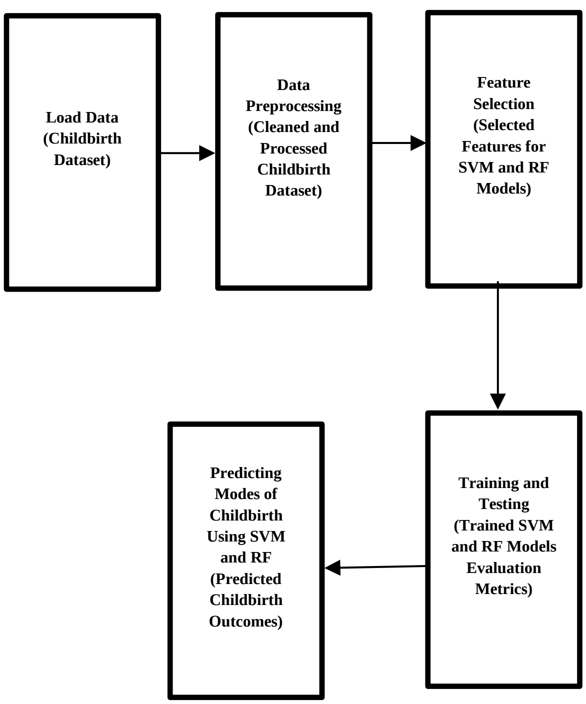

Figure 1.Block diagram

4.2 DATA PREPROCESSING

The dataset is carefully preprocessed in order to improve its fit for machine learning techniques. To guarantee data quality, missing values are handled, outliers are controlled, and any required transformations are carried out. By reducing biases and inconsistencies in the dataset, this stage is essential in creating the foundation for precise and trustworthy predictions.

4.3 FEATURE SELECTION

A crucial step in developing a model is feature selection, which seeks to pinpoint the most important factors affecting the modes of birthing. This stage entails selecting a selection of features that considerably increase the model's predictive power by applying statistical techniques or domain expertise. The final SVM and RF models' accuracy and interpretability are greatly influenced by the characteristics that were chosen. Certainly! Here are two additional points to enhance the paragraph. Furthermore, feature selection aids in reducing model complexity, thereby improving computational efficiency and reducing the risk of overfitting, which occurs when a model performs well on training data but poorly on unseen data.

4.4 TRAINING AND TESTING

Support Vector Machine (SVM) and Random Forest (RF) model training is made possible by the dataset's partitioning into training and testing sets. To produce precise forecasts, these models analyze the data to find patterns and relationships. Strict evaluation measures are used to evaluate the models' performance on the testing set, including as accuracy, precision, recall, F1 score, and AUC-ROC. This stage is crucial for determining how well the selected algorithms predict the outcomes of childbirth.

4.5 PREDICTING MODES OF CHILDBIRTH USING SVM and RF

In order to forecast the modes of birthing, the trained SVM and RF models are really applied in this module. By utilizing the knowledge acquired throughout the feature selection and training stages, the models are able to categories new cases as either vaginal or cesarean deliveries. The predictions produced by these models advance the field of predictive analytics in healthcare by providing useful tools for understanding and forecasting birthing outcomes. Because both SVM and RF are strong algorithms for handling complicated, multidimensional datasets, their combination provides a solid method.

5. ALGORITHM DETAILS

A. Support Vector Machine

Support Vector Machine (SVM) is a popular supervised machine learning model that is used for classification and prediction of unknown data. It is asserted by several researchers that SVM is a very accurate technique for text classification. It is also widely used in sentiment classification. For instance, if we have a dataset in which data is pre-labeled into two categories: positive and negative reviews, then we can train a model to classify new data into these two categories. This is exactly how SVM works. It is the model that we train on a dataset, so it can analyze and classify unknown data into the categories that were present in the training set. SVM is a linear learning method. It finds an optimal hyper-plane to differentiate two classes. Being a supervised classification model, it tries to maximize the distance between the closest training point and either class so as to achieve better classification performance on test data. The process for classification functions is as follows:

- It takes the labeled sample of data, and draws a line separating the two classes. This line is called the decision boundary. The solution is based only on those training data points which are really close to the decision boundary. The data points are called Support Vectors. For example, if we are categorizing movie reviews (in our case), one side of the boundary will have positive reviews while the other side has negative reviews.

- Now when new data needs to be classified, it goes either into the left or right side of the decision boundary. Depending on which side the data enters, it is classified under that category. To classify our data with the best precision, we need to split the two categories such that the decision boundary separates the two classes with maximum space between them.

B. Random Forest Classifier

Random Forest uses hyper parameters similar to Decision Trees and Bagging Classifiers. Combining a decision tree with a bagging classifier is unnecessary when using the random forest's classifier-class. You can use the algorithm's repressor to do regression tasks with random forest. Using a random forest during tree development increases the model's flexibility. When splitting a hub, it selects the optimal component from a diverse collection of supplies, rather than the principal piece. There are numerous alternatives available, with the majority leading to a superior model. The calculation for separating a hub only examines an irregular subset of items in random timberland. To increase the randomness of trees, consider utilizing irregular bounds for each component instead of searching for the optimal edges. The Random Forest approach combines many decision trees into a classification and regression ensemble. A random subset of the training data sets is used to construct multiple decision trees. Furthermore, the ensemble nature of Random Forest reduces the risk of overfitting by averaging the predictions from multiple trees. Using many decision trees improves the accuracy of results. This approach has a short runtime and can handle missing data. Random forest is used to randomize the procedure rather than the training dataset. Decision trees generate decision classes.

1.Here the author firstly shows the Random Forest creation pseudocode: Randomly select “K” features from total “m” features where k << m.

2. Among the “K” features, calculate the node “d” using the best split point.

3. Split the node into daughter nodes using the best split.

4. Repeat the a to c steps until “l” number of nodes has been reached.

5. Build Forest by repeating steps a to d for “n” number times to create “n” number of trees.

In the next stage, with the random forest classifier created, we'll make the prediction. The pseudo code for random forest prediction is shown below:

1. Takes the test features and it uses the rules of each randomly created decision tree to predict the outcome and later stores the predicted outcome (target).

2. Calculate the votes for each predicted target

3. The final prediction comes from high voted predicted target from the random forest algorithm. The process is easy to understand, butit's somehow efficient.

6. RESULT ANALYSIS

In our project, we conducted a comparative analysis of machine learning algorithms to predict birthing modes, revealing varying levels of predictive accuracy. The Decision Tree (DT) algorithm achieved an accuracy of 60%, while K-Nearest Neighbors (KNN) demonstrated 70% accuracy. The Individual Random Forest (RF) model showed improved performance with an accuracy of 80%, and Support Vector Machine (SVM) performed slightly lower at 63% accuracy. Notably, combining the SVM and RF algorithms resulted in a significant enhancement, yielding the highest accuracy of 88%. This synergistic approach effectively leveraged the strengths of both SVM and RF, highlighting their complementary nature in improving predictive accuracy. The integration of these algorithms demonstrates promising potential for enhancing birth outcome predictions, emphasizing the value of ensemble methods in medical decision-making environments.

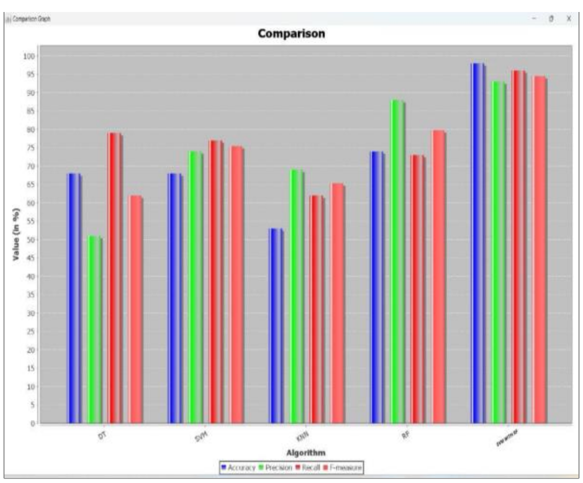

Figure 2. Comparison graph

7. CONCLUSION

In conclusion, our study underscores the potential of combining Support Vector Machines (SVMs) and Random Forests (RFs) to enhance the accuracy of predicting the mode of delivery. This approach offers a promising avenue for achieving more precise and cost-effective birthing outcomes, reducing unnecessary cesarean sections, and mitigating associated risks for both mothers and newborns. Additionally, the enhanced predictive accuracy can empower expectant mothers with valuable insights into their birthing options, fostering shared decision-making and personalized care plans tailored to individual needs and preferences. Furthermore, the integration of advanced machine learning techniques like SVMs and RFs can pave the way for continuous learning and improvement, as the system evolves with the accumulation of more comprehensive and diverse datasets. This iterative approach can contribute to ongoing advancements in obstetric care, ultimately leading to better maternal and neonatal outcomes and enhancing overall healthcare quality and efficiency.

8. FUTURE WORK

Future research endeavors will focus on expanding the dataset to encompass a broader and more diverse range of expectant mothers, thereby enhancing the system's accuracy and generalizability. By incorporating a more extensive dataset, the machine learning models can be trained to recognize and adapt to a wider spectrum of maternal and fetal variables, further refining the predictive capabilities of the system. This would boost patients' confidence in the system's predictions and assist clinicians in understanding why the system produces the predictions that it does. Moreover, future research will explore the integration of real-time data streams and continuous monitoring capabilities to enable dynamic updates and adjustments to the predictive models.

9. REFERENCES

1.In Mishra, S. Cesarean delivery prediction with ensemble machine learning techniques. Electronics 12, 3524 (2023).

2. Alberto De Ramón Fernández, Daniel Ruiz Fernández. Prediction of the mode of delivery using artificial intelligence algorithms. 219 (2022) 106740

3. Muhammad nazrul islam, Tahasin mahmud, Nafiz imtiaz khan, Sumaiya nuha mustafina. Exploring Machine Learning Algorithms to Find the Best Features for Predicting Modes of Childbirth. 2020.3045469

4.Investing the prediction of emergency cesarean section by measurable maternal and fetal characteristics.Ping Guan, Fei Tang, Guoqiang Sun, Wei Ren. 2020;68;799–806.

5. Ms. Michal Lipschuetz, RN MPH MSc, Mr. Joshua Guedalia, MBA, MD, Michal Novoselsky Persky. Prediction of vaginal birth after cesarean deliveries using machine learning. YMOB 13033.

6. Prediction of Cesarean Childbirth using Ensemble Machine Learning Methods. Nafiz Imtiaz Khan ,Tahasin Mahmud, Muhammad Nazrul Islam. 978-1-4503-8922-8/20/11.

7.Dorri, A., Jurdak, R., Kanhere, S.S., and Gauravaraam, P. Lsb: investigating machine learning algorithms to determine which features are most useful in predicting birthing styles. A. J. Separate distribution. Computer. 134, 180–197, 2019.

8.Kumar, n.; Mistry, i.; Tanwar, s.; Tyagi, s. A comparison of the decision tree, random forest, and xgboost models' performances is shown, along with an interpretation of the results using Shap for determining when a caesarean section of labor is necessary. Mech. System. process of signals. 2020, 135–106 38382

9.Sabeen, S.; Felcia, H.J. Systems that indicate the best way to give birth are based on machine learning. 2022, 19, 3741–3763 in Webology

10.Maamar, z.; faci, n.; ugljanin, e.; baker, t.; burégio, v.: Using a machine learning algorithm to predict the best delivery method 2021, 14, 100400.

11. a. narayanan; n. feamster; huang, d.y.; reisman, d.; apthorpe, n. application of artificial intelligence techniques to predict the delivery modality. Proceed. Exclusive. improving technology. 2020, 2019, pages 128–148.

12.Pal et al. (year): Preferred and actual mode of delivery in relation to dread of childbirth j.; Rabehaja t.; Hill a.; Hitchens m.; Varadharajan v. 7, 2630–2639, 2020.

13.Qu (c.), Tao (m.), Zhang (j.), Hong (x.), and Yuan (r.). Scots population data is used to generate individual prediction ratings for the mode of delivery following a cesarean section. Connect. Webw. 7817614, 2018, 2018, 2018.

14.Ariaratnam, S.T.; Alnahari, M.S. A case study of Bangladesh utilizing bagging ensemble classifier for birth mode prediction, 5, 979–993.

15.Rehmani, M.H.; Chen, J.; Hassan, M.U. Classification of childbirth techniques using machine learning models and feature selection. Future gener. Computer. System. 2019, 97(5), 524–529.