



Predictive Analytics for Childbirth Mode Classification Using Machine Learning Techniques

Marpu Pradeepthi¹, Asoda Manasa²

Assistant Professor, Department of CSE, Megha Institute of Engineering and Technology for Women,
Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana – 501301¹

M.Tech Student, Department of CSE, Megha Institute of Engineering and Technology FOR Women,
Edulabad (Village), Ghatkesar (Mandal), Medchal District, Telangana – 501301²

Abstract: The birthing process is crucial to the mother's health and the relationship she forms with her infant. For the sake of both mother and child, the choice about the mode of delivery must be swiftly implemented. It is challenging for healthcare providers to make prompt and accurate choices in this area since mistakes in this area may have major effects on the mother's and fetus's health. A machine learning-based decision-support system for determining the most secure delivery mode is introduced in this study. K-Nearest Neighbours (KNN), Random Forest (RF), Decision Tree, Support Vector Machine (SVM), and a stochastic classifier are among the supervised learning techniques that are evaluated. This paper's results show that the Random Forest algorithm is more accurate at predicting the delivery mode, which helps doctors make better decisions and keeps mums and babies safer.

Keywords: childbirth prediction, machine learning, decision support system, delivery mode classification, maternal and infant health

INTRODUCTION

In high-risk situations requiring quick decisions, the technique of delivery determines the mother's and the unborn child's well-being. There are a lot of medical factors to think about when deciding whether to have a natural or caesarean birth, and decisions may easily go wrong, causing serious problems. Under time constraints, the traditional method of obstetrical decision-making—which heavily depends on the physician's expertise—often leads to less-than-ideal results. To mitigate these dangers and improve the result, clinical decision-making using machine learning may be a good solution. The authors of this research recommend that doctors and nurses use a decision-support system based on machine learning to determine the most secure delivery method. The system's foundation is built on supervised learning algorithms that analyse patient data and provide suggestions in real-time. These algorithms include Stochastic Classifier, Random Forest, Support Vector Machine, and Decision Tree. The system determines patterns and the most important elements in determining delivery outcomes by learning from previous birth data, which aids clinical decision making. Reducing the possibility of inaccurate mode of delivery selection, improving safety for women and infants, and assisting physicians in making optimum judgements within the given time limit are the main aims of this research. By quickly calculating from massive amounts of patient data, the proposed technology provides a certain instrument to enhance the accuracy and efficiency of medical choices connected to birthing.

PREDICTION OF BIRTH

"Childbirth prediction" refers to the practice of assessing a number of clinical factors—such as the mother's age, health history, foetal position, and pregnancy concerns—to ascertain whether a caesarean section or vaginal birth would be the safest option. Aiming to anticipate possible challenges and allow proactive preparation, accurate childbirth prediction aims to decrease the need for emergency interventions. By using machine learning models, this software enhances the accuracy of predictions, giving obstetricians a data-backed way to make important choices during labour and delivery. As a result of better prediction accuracy, outcomes are improved and the primary goal of improving maternal care public health is advanced, leading to lower risks for the mother and the foetus



ELECTRONICS

Machine learning is an artificial intelligence (AI) approach to improving computers' automated performance on certain tasks. It aims to do this by replacing data-driven learning with explicit programming. Reason being, it enhances operational effectiveness and forecasting abilities by transforming data into practical insights. Machine learning algorithms analyse large amounts of data related to delivery outcomes, finding complex patterns and correlations that aid healthcare providers in making quick, educated choices. In the long run, this predictive capacity helps make healthcare safer and more personalised by cutting down on human mistake, increasing clinical precision, and tailoring treatment plans to each patient's unique needs.

System to aid in decision-making

Electronic health records, clinical guidelines, and prediction models are just a few of the many sources of information that a Decision Support System (DSS) compiles to aid doctors in making choices. Individualised delivery is one feature offered by the DSS in this research. techniques via the integration of machine learning predictions with real-time patient data. This technology improves the efficacy and efficiency of healthcare systems by facilitating faster response times in high-pressure situations and helping to standardise decisions based on best practices and data-driven insights. Quick and precise decision-making is facilitated by DSSs, leading to fewer difficulties and more patient satisfaction. Additionally, they support efforts to create precision medicine programmes for healthcare providers.

Classification of Delivery Modes

The term "delivery mode classification" describes the process of categorising various methods of giving birth into different categories, such as caesarean section and vaginal delivery. During this categorisation process, we look at the patient's present physiological data, their medical history, and any signs that may be particular to pregnancy. In order to improve the reliability of assessments, this research makes use of ML modalities such support vector machines and decision trees to differentiate between different delivery methods. Medical personnel may better ensure the safety of both mother and child by classifying delivery strategies according to the risks associated with each and then choosing the method that poses the least risk.

LITERATURE SURVEY

One objective put out by Eri Maeda and colleagues [1] in their article is: Japan lacks comprehensive descriptive data on caesarean deliveries due to the absence of national prenatal registries. We intend to characterise caesarean section births across single and multiple pregnancies using the Japanese National Database, which aggregates health insurance claims. The national and prefectural rates of caesarean deliveries for 2014 were calculated proportionately for both single and multiple pregnancies. There was talk on where and what kinds of prenatal and newborn care facilities are available, as well as about blood transfusions and other maternal morbidities. While 18.6% of all births were caesarean sections, 82.7% of women who had multiple pregnancies did so. The prefectural rates of caesarean births for total pregnancies ($r = 49.2\%$ to 100% and $r = 12.5\%$ to 24.2% , respectively) and multiple pregnancies ($r = 0.59$, $p < 0.001$) were generally well correlated. Overall, 1.4% of caesarean patients had allogeneic blood transfusions, whereas 3.2% of patients with multiple pregnancies received such transfusions. Qian Chen and coworkers [2] have suggested in this study Evidence that a caesarean section may have long-term effects on the baby's health is mounting. The underlying mechanisms, however, are still a mystery. This research seeks to answer the question of whether a caesarean section done at the mother's desire and without medical reasons (CDMR) affects the DNA methylation state of the infant's umbilical cord blood. A cross-sectional research was conducted in Shanghai, China. It was in 2012 when 70 CDMR and 70 VD were recruited. A total of thirty-five CDMR and thirty VD newborns' cord blood samples were tested for DNA methylation using the Illumina Infinium Human Methylation 450 K Bead Chip. To validate the results, the DNA methylation status in the cord blood of 40 extra CDMR and 40 VD newborns was evaluated using a targeted bisulfite sequencing method. The research comprised 40 genes with a total of 497 CpG sites. It is unclear if early life risk factors contribute to childhood obesity in preterm neonates; little is known about the effects of eating habits, and this was highlighted by Yuanqing Fu and colleagues [3]. We looked for dietary patterns that might change these risk factors, as well as early warning signs of childhood obesity and overweight in premature infants. There are 338,413 mother-child pairs in the Jiaying Birth Cohort (1999–2013), with 2125 being suitable singleton preterm infants. We took anthropometric measurements, dietary and lifestyle data, and medical assessments from each participant while they were in the clinic.

Using a machine learning-based interpretable analytical technique, we identified life-time predictors of childhood obesity and overweight, and we used Poisson regression to look for connections between eating habits and the most



relevant predictor. Among the 2125 eligible preterm children, 274 (12.9 percent) (863, or 40.6%) were overweight or obese between the ages of 4 and 7. You, Muhammad. Using the most effective features for determining the delivery mode (vacuum extraction, forceps delivery, emergency caesarean, vaginal birth, or caesarean birth), the study aimed to evaluate machine learning algorithms and identify viable characteristics for determining the delivery mode. An empirical enquiry was conducted using a literature review, interviews, and a structured survey to effectively identify the most essential machine learning approach for prediction using 6157 birth data and a standard assortment of characteristics. This research by Abin Abraham et al. [5] suggests that prenatal care may be improved if we knew which pregnancies were likely to have preterm births, which is a leading cause of infant mortality worldwide. Unfortunately, there aren't any popular methods that can accurately predict the likelihood of a preterm birth just yet. Despite the fact that EHRs allow for the scalable and inexpensive estimation of illness risk, the vast longitudinal data accessible in EHRs has not been well used in pregnancy research. In this study, we forecast preterm delivery in singletons by using machine learning to several electronic health record datasets. Using a massive dataset consisting of 35,282 newborns, the researchers found that a billing code-only prediction model outperforms a similar model trained using known risk factors (ROC-AUC=0.59, PRAUC=0.21) and can identify preterm delivery as early as 28 weeks gestation (ROC-AUC=0.75, PRAUC=0.40).

Innocent B. Mboya et al. proposed comparing machine-learning methods to the logistic regression model in this study to identify the main factors contributing to newborn fatalities. This research used secondary data from the Medical Birth Registry cohort at Kilimanjaro Christian Medical Centre (KCMC) from 2000 to 2015. We evaluated the discriminatory power of the models using decision curve analysis and the area under the receiver operating characteristics curve (AUC). Located in Moshi Municipality in northern Tanzania, the Kilimanjaro area is home to the KCMC, a zonal referral hospital. On the premises of the hospital is housed the Medical Birth Registry of the Reproductive and Child Health Centre. Count every singleton birth that occurred between the years 2000 and 2015 ($n = 42,319$). Stillbirths and early neonatal deaths are also included in the perinatal death category. Prior to the mothers' discharge from the hospital, these outcomes were only recorded. There were 3.7 percent of maternal fatalities. Research by Mohammad Khubeb Siddiqui et al. indicates that neuronal signals may be used to diagnose epilepsy, a serious and chronic neurological disorder. Neurones are highly interconnected so that they may form messages and interact with other parts of the body. It is common practice to detect these brain impulses using electrocorticography (ECoG) and electroencephalography (EEG) media. These signals generate a great deal of data while also being complicated, noisy, non-linear, and non-stationary. Therefore, it is not an easy task to learn about the brain or to recognise seizures. Seizure detection, EEG data classification, and the identification of pertinent, logical patterns are all tasks that ML classifiers can handle. As a result, many researchers have created tools—including statistical characteristics and machine learning classifiers—for seizure identification. The biggest obstacle is picking the right classifiers and features. Different "black-box" and "non-black box" statistical characteristics and ML classifiers have been the basis for these techniques in recent years, and this work intends to summarise all of them. Research by Yan Xi SOH and colleagues has proposed This research aims to use structural equation modelling to probe the connections between obstetric and sociodemographic variables, psychological health, fear of giving birth, and self-efficacy during delivery. The research, which included 205 pregnant women from diverse ethnic backgrounds in Singapore, was cross-sectional and exploratory in nature. Using structural equation modelling, we tested a theoretical model that incorporates Bandura's self-efficacy theory with findings from previous literature studies. Psychological well-being was assessed using the World Health Organisation (Five) Well-Being Index; fear of delivery was measured using the delivery Attitudes Questionnaire; and birthing self-efficacy was assessed using the Birthing Self-Efficacy Inventory. According to Yanfang Chen et al., there has been contradictory data about the connection between delivery technique and PTSD in previous studies. In attempt to better understand the connection between the mode of birth and post-traumatic stress disorder (PTSD), this research examined a sample of Chinese women who had a high number of caesarean deliveries. From 2019–2020, researchers in Guangdong, China, followed a cohort of participants from October to August. After obtaining written permission, we contacted and recruited pregnant women (aged 20–45) who had previously received prenatal care at Nanhai Hospital, which is part of Southern Medical University, and who intended to give birth there during the study period. The PTSD Checklist—Civilian Version was used for the purpose of evaluating chronic stress disorder. The obstetric and sociodemographic characteristics of the groups who had vaginal and caesarean births were first compared. According to a study by Yiye Zhang et al., there are just a few methods that may be used to predict postpartum depression (PPD). We provide a machine learning approach to PPD risk assessment that makes use of data extracted from EHRs. Two electronic health record datasets were used to construct the PPD risk prediction model. One dataset included information on 15,197 women from a single location between 2015 and 2018, while the other included information on 53,972 women from various sites between 2004 and 2017. A PPD diagnosis was the primary outcome within one year after giving birth. Using a method that included data extraction, processing, and machine learning, a core set of features was selected from the EHR datasets to ensure model performance and enable future point-of-care risk prediction. The top-performing model incorporates clinical data including patient information, obstetric difficulties, medical comorbidities, prescription orders for treatment, and prescriptions for medications. Tasks Connected



How content a mother is with her delivery may have short- and long-term effects on her health and bond with her child. What happens during delivery has a profound impact on the mother and her baby. It could be a crucial component for the mother's and child's protection. It becomes very difficult for the doctor to make rapid judgements during a baby's birth. It is also possible for people to pick the wrong delivery method when giving birth. Making a bad decision may raise the mother's risk of death and harm the health of the baby. One promising approach to this challenge would be the use of computers to make decisions. In order to lessen the impact of this danger, we used supervised machine learning to determine the optimal distribution approach and then developed a decision-making system around it. The real birth information from Kisorj, Bangladesh's Tarail Upazilla Health Complex were fed into 32 supervised classification algorithms and 11 training techniques. We have also compared the outcomes using several statistical criteria in an effort to determine the most effective model.

METHODOLOGY

We want to develop a decision support tool for medical professionals to use in predicting the optimal delivery strategy using the proposed technology. This system incorporates many supervised learning approaches, such as Random Forest, k-Nearest Neighbours (KNN), Support Vector Machine (SVM), and Stochastic Classifier, due to their specific benefits in classification tasks. The system is programmed to analyse patient data, such as the mother's medical history, the foetal state, and other relevant medical features, in order to accurately predict whether a vaginal birth or a surgery is preferable. The technology can determine which delivery technique is most appropriate by analysing patterns and factors in previous birthing data used to train the models.

The first step is to load the dataset.

In the first module, the dataset is collected and loaded. It contains medical records and factors that are crucial to birthing decisions. Collecting patient vitals such blood pressure, history of pregnancies, mother's age, and any issues with the foetus is standard procedure. The information may also contain labels for previous ways of delivery, such as the possibility of a vaginal birth or a caesarean section. It is critical to load the dataset properly and rapidly since it provides the basis for the whole decision-making process.

Preprocessing of Data

Typical input dataset artefacts like as noise, outliers, missing values, and so on might have an effect on the model's accuracy. To make it usable by machine learning algorithms, this raw data has to be cleaned up and organised by the data preparation module. Missing Value Management: As part of the preparation methods, missing values are either imputed or deleted based on their relevance and the overall quality of the dataset. To make the model training process more efficient, continuous variables such as weight, age, and blood pressure are normalised or standardised to a single scale. Algorithms can handle categorical data (such as delivery history and maternal health status) by converting it to numerical values using techniques like label encoding and one-hot encoding. Finding and removing outliers that can skew the model's learning process helps maintain the dataset representative of real-world scenarios. By ensuring that training and testing are carried out on high-quality data, effective data preparation enhances the accuracy and reliability of machine learning models.

Feature Selection

The main goal of this module is to identify the dataset characteristics that significantly affect the delivery mode prediction. Feature selection improves the dataset's performance by reducing the dimensionality by removing characteristics that aren't needed or duplicated. A number of techniques are used, such as statistical testing, correlation analysis, and recursive feature elimination (RFE), to determine which qualities are most strongly associated with the dependent variable (delivery mode). Some maternal health problems, such hypertension or gestational diabetes, may have a greater impact than other variables, like as the mother's previous delivery experience.

Education and Assessment

The machine learning method revolves on this module, which divides datasets into training and testing sets. Machine learning makes it easy to spot relationships and patterns in data; the testing set is used to assess the model's performance on anonymous data, while the training set is used to educate the model. To understand how different qualities interact and impact the choice of delivery method, machine learning models such as SVM, Random Forest, Decision Tree, KNN, and Stochastic Classifier are trained using the provided data. Training aids the system in comprehending the complex relationships between medical issues and delivery outcomes; every algorithm has several uses. At this stage, you'll tweak the hyper settings to get the most out of the model. After training, the model is put to the test using the testing set to see how well it can generalise to new, unknown occurrences.



Categorisation

This lesson concludes with a classification and forecasting exercise using the learned machine learning model for optimal delivery. When new patient data is submitted, the system examines input factors, such as foetal characteristics and maternal health problems. It then makes a prediction on the best course of action, whether it a spontaneous birth or a caesarean section. In order to enable healthcare professionals make a faster and more informed decision about the delivery method, the classification results are provided as recommendations. The classification module facilitates faster decision-making during labour, which is particularly useful in high-risk or emergency situations. It reduces the likelihood of unsavoury outcomes that could emerge from poor or delayed decision-making and enhances patient safety by providing accurate, data-driven insights.

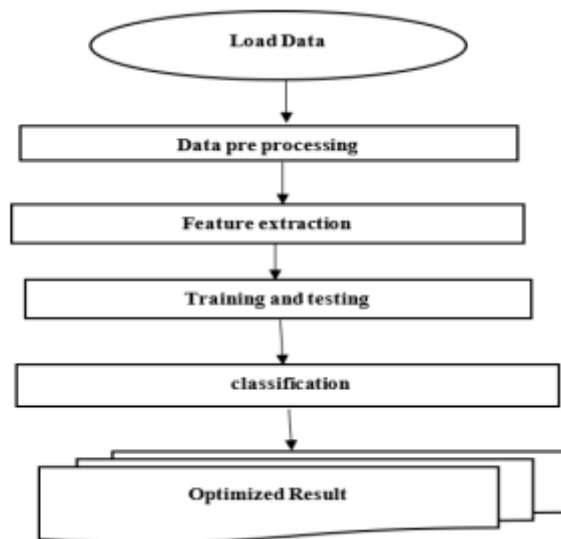


Fig. 1. Block diagram

RESULT ANALYSIS

We use a variety of metrics to assess the efficacy of the applied algorithms, including Support Vector Machine (SVM), Random Forest, Decision Tree, k-Nearest Neighbours (KNN), and Stochastic Classifier, in order to determine the best method of delivery for an ML-based decision support system. Researchers successfully evaluated the models' F1-score, recall, accuracy, and precision using an autonomous dataset to completely comprehend their prediction potential. Finding intricate patterns in the data while avoiding overfitting allowed the Random Forest technique to consistently outperform the other models in terms of accuracy, as shown in the research. It seems that SVM is effective in lowering false positives, as it showed promising results, particularly in accuracy. Although the Decision Tree model is simple to grasp, it is not very accurate since it is easily fooled by noisy input. Depending on the distance measure and value of k, the KNN algorithm's performance was average at best. Although it was not widely used, the Stochastic Classifier provided useful information about the uncertainty of predictions. The data analysis demonstrates that the proposed method has the potential to improve clinical judgement, be of use to women and babies, and provide accurate predictions for the categorisation of delivery types. It is possible that this system may become an invaluable resource for doctors with the addition of more data and tweaks to its hyperparameters.

algorithm	accuracy
EXISTING	75
PROPOSED	88

Fig. 2. Comparison Table

CONCLUSION

Finally, by developing an ML-based decision support system to choose the best delivery method, both mother and child health may be significantly enhanced. To efficiently evaluate patient data and provide precise and prompt delivery method



predictions, the system makes use of a variety of supervised learning techniques, including k-Nearest Neighbours, Random Forest, Decision Tree, Support Vector Machine, and Stochastic Classifier. The findings analysis demonstrates how these models have the potential to enhance the safety of both the mother and the infant during delivery by reducing the likelihood of human mistake in critical decision-making contexts. The importance of incorporating technology into healthcare and the need to continually refine and adapt models to different clinical settings are both emphasised by this research. Further study and validation of this decision support system may make it an indispensable resource for healthcare practitioners, leading to safer delivery processes and better decision-making.

FUTURE WORK

In order to make the decision support system that uses machine learning more effective and acceptable for use in clinical settings, future research will concentrate on a few important aspects. The first and most important step is to expand the dataset to include a broader variety of patient demographics, medical histories, and delivery complications. This will improve the model's accuracy and make it more applicable to various demographics. Looking at state-of-the-art ML techniques, such as deep learning and ensemble methods, improves the system's prediction ability even more. The integration of the decision support system with EHRs, which allow for real-time decision-making and data accessibility in clinical settings, should also be the focus of future study. Conducting prospective studies to validate the system's recommendations in real-world contexts is necessary to demonstrate its reliability and inspire trust among healthcare providers.

REFERENCES

- [1]. Eri Maeda, et al., "Cesarean delivery rates for overall and multiple pregnancies in Japan: A descriptive study using nationwide health insurance claims data." *Journal of Obstetrics and Gynecology Research* (2021).
- [2]. Chen, Qian, and colleagues' article "The impact of cesarean delivery on infant DNA methylation" appeared in *BMC Pregnancy and Childbirth* 21.1 (2021): 1–8.
- [3]. Fu, Yuanqing, and their fellow researchers, "Integration of an interpretable machine learning algorithm to identify early life risk factors of childhood obesity among preterm infants: a prospective birth cohort." *BMC Medicine* 18.1 (2020), 1–10.
- [4]. Exploring Machine Learning Algorithms to Find the Best Features for Predicting Modes of Childbirth," by Muhammad Nazrul Islam et al., *IEEE Access* (2020).
- [5]. Dense phenotyping from electronic health records enables machine learning-based prediction of preterm birth," Abraham, Abin, and associates. *medRxiv* (2020).
- [6]. Mboya, I.B., et al., "Prediction of perinatal death using machine learning models: a birth registry-based cohort study in northern Tanzania," *BMJ Open*, vol. 10, no. 10, e040132, 2020.
- [7]. Siddiqui, M.K., et al., "A review of epileptic seizure detection using machine learning classifiers," *Brain Informatics*, vol. 7, pp. 1–18, 2020.
- [8]. Soh, S.E., et al., "Determinants of childbirth self-efficacy among multiethnic pregnant women in Singapore: a structural equation modeling approach," *Midwifery*, vol. 87, p. 102716, 2020.
- [9]. Chen, Y., et al., "Association between mode of birth and posttraumatic stress disorder following childbirth: a prospective cohort study of Chinese women," 2020.
- [10]. Zhang, Y., et al., "Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women," *Affective Disorders Journal*, vol. 279, pp. 1–8, 2020.