

Exploring Machine Learning Algorithms to Find the Best

Features for Predicting Modes of Childbirth

Mr.K.Vijaya Babu ¹, Amal Thomas², Vengala Aksshaya², Kalvala Sneha²,
Vanaparathi Vamshi Varma^{2,1Asst} Professor, Computer Science and Engineering,
CMR Engineering College, medchal, T.S, India² B.Tech, Computer Science and
Engineering, CMR Engineering College, medchal, T.S, India

Abstract: A key factor in assuring the safety of both mother and child is the delivery method. Choosing the incorrect delivery style may result in many short- and long-term health complications for both the mother and the infant. However, the current standard of care for predicting the mode of birth usually relies on the judgement of the attending physician. The aim of this study was twofold: first, to identify potential features for predicting the mode of childbirth (vaginal birth, caesarean birth, emergency caesarean, vacuum extraction, or forceps delivery); and second, to investigate machine learning algorithms using the best features for doing so. Five different machine learning algorithms were investigated in an empirical study that used 6157 birth records and a minimum set of features to explore the relevant features for predicting the mode of childbirth. This study also included a literature review, interviews, and a structured survey. The study identified 32 variables that may be used to predict different delivery methods and grouped the data into several categories according to their significance. When all (n D 32) features were taken into account, stacking classification (SC) produced the highest f1 score (97.9%), followed by random forest (RF), decision tree (DT), and support vector machine (SVM), which both performed nearly as well (f1-score D 97.3%).

1. INTRODUCTION

One or more infants exit the womb at the end of pregnancy either naturally or by a caesarean section. Natural childbirth is generally the most frequent birthing method, followed by caesarean sections, emergency caesarean sections, vacuum extractions, and forceps deliveries [1]. Each of these birthing scenarios has advantages and disadvantages, and the style of delivery selected may not be appropriate given the mother's personality. The improper delivery method may result in a variety of hazards, including foetal termination, severe bleeding, the newborn having respiratory difficulties, and more [2]. Even though natural childbirth is the most frequent method, it may be complicated for symptomatic moms, such as those who are older than 35, have conditions like diabetes or preeclampsia, or are carrying

Mr.K.Vijaya Babu *et. al.*, / International Journal of Engineering & Science Research

multiple foetuses [3]–[6]. When a regular birth might endanger the mother or the baby, a caesarean section, sometimes referred to as a c-section or caesarean delivery, is frequently required [7]. Even while a caesarean section may not be the ideal option to a normal birth, it may be a lifesaving treatment for both mother and child in some situations, such as when a baby is in an unusual position in the womb or labour is not proceeding as it should.

In addition to these issues, the prevalence of caesarean sections is rising worldwide. The World Health Organisation (WHO) reports that between 2000 and 2015, the percentage of births that included a caesarean section almost quadrupled, rising from 12% to 21% [8]. The total rate of maternal mortality ranged from 6 to 22 per 100,000 live births, with between one-third and fifty percent of maternal fatalities occurring after caesarean delivery [9]–[11]. In poor and undeveloped nations, the situation is worse; for instance, in Bangladesh, the incidence of needless caesarean procedures rose by 51% between 2016 and 2018 [12] and 77% of all caesarean operations in 2018 were medically unnecessary. The maternal mortality rate in Bangladesh in 2017 was 173 deaths per 100,000 live births, which was greater than the rate in the majority of affluent nations; in the same year, the rate in the United States was 17 fatalities per 100,000 live births [13]. Furthermore, the National Low Birth Weight Survey Bangladesh, 2015 conducted by the Institute of Public Health Nutrition revealed a caesarean section rate of 35.5%, which was higher above the WHO-recommended range of 10-15% [14], [15]. Maternal mortality and morbidity rise to around double the rate for caesarean birth as compared to normal delivery. Mothers who have caesarean sections may experience a variety of health problems, including blood loss, organ damage, infection, trouble with subsequent pregnancies, and similar concerns. [16]. The surgical caesarean section surgery, often known as an emergency caesarean section, may sometimes be urgently required owing to acute concerns for the health of the mother and/or infant [17], [18]. Emergency caesarean sections had a greater risk of surgical damage and infection while being identical to traditional caesarean sections [19].

A Description Of The Project: Therefore, the goals of this study were to (a) investigate and rank the features that are most crucial for predicting delivery modes, and (b) investigate machine learning techniques and create a number of frameworks, based on various machine learning algorithms, for predicting delivery modes reasonably accurately with a small number of features. A literature research, interviews, and a short survey were conducted to investigate and rank the necessary elements in order to accomplish these goals. In order to determine the

most efficient delivery methods, a number of prediction models based on decision tree (DT), k-nearest neighbours (KNN), random forest (RF), support vector machine (SVM), and stacking classification (SC) techniques were proposed. The performance of these models was assessed in terms of precision, recall, and f1 scores.

2. LITERATURE SURVEY

2.1 Existing System

By identifying obstetric risk variables via data mining, Pereira and colleagues [20] performed a research to predict the kinds of deliveries (normal, caesarean, forceps, and hoover). They used DT, GLMs, SVMs, NB, and other data mining methods in their work. Among these models, DT produced the most positive outcomes for statistical indicators, with the greatest accuracy and specificity. Some research compared Bishop's scores and translabial ultrasound measures to assess the appropriateness of inducing labour. For instance, Khazardoost and colleagues [21-22] examined Bishop's scores and translabial ultrasound measurements to assess the values of ultrasound measurements relevant to the outcomes of births. Translabial ultrasonography was used to assess the foetal head pubis symphysis distance and the cervical length. Multivariate analysis was used to establish the predictive value of the Bishop's score, cervical length, and foetal head pubis symphysis distance. The findings demonstrated that translabial measures were a more effective way to track the development of labour than the Bishop's score.

Ramanathan and colleagues [23] looked at the possible consequences of labour, such as the possibility of a protracted pregnancy and the need for a caesarean section as a result of an unsuccessful induction. According to the study's findings, measuring the cervical length at 37 weeks might predict the likelihood of needing an emergency caesarean surgery. Premature birth risk, vaginal delivery after a caesarean section, and other dangers associated with pregnancy have been the subject of certain research. For instance, a research by Lipschuetz and colleagues [24] used gradient boosting, RF, balanced RF, and AdaBoost ensembles to create a customised tool for forecasting vaginal births after caesarean deliveries.

Similar to this, Tessmer-Tuck and colleagues [25] created a model utilising multivariate analysis to predict vaginal births after caesarean procedures. The multivariate model constructed using characteristics provided by the Grobman prediction model [26] had an area under the curve of 75.7%, while the multivariate model utilising features assessed using

Mr.K.Vijaya Babu *et. al.*, / *International Journal of Engineering & Science Research* stepwise regression had an area under the curve of 72.3%. In a different research, Brando and colleagues [27] used DT, SVMs, and GLMs to conduct classification in order to assess the voluntary termination of pregnancy and determine the associated risks for patients. The study's SVMs had an accuracy of 68% and a sensitivity of around 93%. Similar to this, Lakshmi and colleagues suggested a prediction model based on a C4.5 classification tree to assess the significance of various prenatal traits or factors for anticipating risk levels and difficulties throughout pregnancy. A decision support system was created by Goodwin and Maher in a different research to precisely identify women who were at risk of preterm delivery and the factors contributing to it.

2.2 Proposed System

Typically, the responsible medical expert chooses the delivery method. To help clinicians forecast a delivery method that is consistent with the mother's features and lessen difficulties during labour, a maternal healthcare application would be beneficial. Even fewer clinical decision-support systems have been created in accordance with such research, and Pereira and colleagues [20] conducted a study to design data mining classification models to predict types of delivery using obstetric risk factors in real time. Few studies have considered the prediction of pregnancy outcomes and modes of delivery.

In a research using ultrasound technology, Usman and colleagues [21] created the "Intrapartum" Android app, which enables medical practitioners to gauge the chances of a normal birth. This system reviews more research. The elements that are essential for forecasting delivery modes must be explored in more detail, and it must be determined which supporting factors might boost the precision of such predictions. Similar to this, research should identify whether algorithms or methodologies may give more accuracy when taking into account a certain collection of traits, since it is uncommon to be able to analyse all of a pregnancy's attributes before or during birthing.

2.2 Shown Features and Their Priorities

The weighted average score from the structured survey was used to identify the first set of features, based on 111 elements collected from the literature survey and structured interviews, in order to find the best collection of features that should be taken into account for predicting delivery procedures. A weighted average score over 1.5 was achieved by 32 features, while a score below 1.5 was achieved by 79 features. 32 characteristics were chosen

Mr.K.Vijaya Babu *et. al.*, / International Journal of Engineering & Science Research

as the first set of features by establishing a threshold value of 1.5, and the other 79 features were deemed to be of considerably lesser importance. The univariate feature selection (UFE) approach was used to rank the original collection of 32 characteristics according to how well they correlated with the goal feature (type of birth) in order to get the best subset of features for prediction. Each characteristic is examined separately by UFE to see how it relates to the dependent variable To determine the prioritised list of characteristics in this research, a chi-squared test was used to categorise the features into high-, medium-, and low-importance categories. There are several additional cutting-edge feature selection techniques that might have been employed to classify the features, including mutual information-based feature selection and LDA-based dimensionality reduction



Fig 1: System Flow

3. SYSTEM ANALYSIS&DESIGN

3.1 DECISION TREE

For this classification task, DT was applied to preset classes of features; the f1 scores for the DTs for all classes are shown in Figure 3 alongside the other four algorithms. DT is a non-parametric supervised learning approach used for classification and regression Table 5 shows that the accuracy, recall, and f1 scores for class five, which had 21 characteristics, were 92.3%, 92.4%, and 92.3%, respectively. For classes four through seven, it can be observed that acceptable values for the accuracy, recall, and f1 scores were obtained. Figure 3

and Table 5 make it evident that class seven, which included all 32 characteristics, had the greatest f1 score (93.2%) and that the accuracy and recall obtained using the DT method were both 93.2%. Precision, recall, and f1 scores totaling all 11 high-importance traits were 82.9%, 82.6%, and 82.8%, respectively. Last but not least, it became clear that the medium-importance feature class outperformed the low-importance feature class while the high-importance feature class outperformed both.

B. UNUSUAL FOREST

For classification, regression, and other tasks, RF is an ensemble learning technique that builds a large number of decision trees during the training phase and outputs the class that represents the mean of the classes (classification) or mean prediction (regression) of the individual trees Figure 3 and Table 5 show the outcomes of applying the RF algorithm to predetermined classes of characteristics. The high-importance class had the greatest f1 score (83.9%) among the low, medium, and high-importance feature classes, while the medium-importance class outperformed the low-importance class. Nevertheless, with the exception of class six, the f1 scores and other performance measures rose from class four forward. Class seven, however, had the strongest performance measure, comprising of all 32 characteristics with accuracy, recall, and f1 scores of 97.3%.

3.2 NEAREST NEIGHBOURS FOR C. K

KNN is a non-parametric, lazy learning method that use a database of data points divided into several classes to forecast the classification of a fresh sample point Figure 3 and Table 5 show the results obtained when using the KNN method to create prediction models based on predetermined classes of characteristics. It is evident that the high-importance class outperformed the medium-importance class, which outperformed the low-importance class in turn. From class four to class seven, the performance of the other classes steadily improved. However, after using KNN, class seven demonstrated the best performance, with precision, recall, and f1 scores of 95.9%, 96%, and 95.8%, respectively.

3.3 System Architecture:SUPPORT VECTORMACHINE

When compared to other machine learning methods, SVM implementation is exceptional since it can handle several continuous and categorical variables SVM was taken into account in the study, and graphs showing the algorithm's performance are shown in Figure 3.

In terms of SVM, it can be shown that performance and important level have a positive association (high, medium, and low importance classes, respectively, had high, medium, and

low performance scores). Although class seven had the greatest accuracy, recall, and f1 scores, SVM's performance grew linearly from class four to class seven. Despite this, SVM performed rather poorly when compared to the other methods for all seven classes of features.

Class Seven had the highest f1 score (88.6%), which included 32 features. Class Six had a f1 score of 85.1%, which included 21 features, and Class Five had a f1 score of 83.4%, which included the same amount of features. Therefore, it may be claimed that class seven had the best SVM outcome.

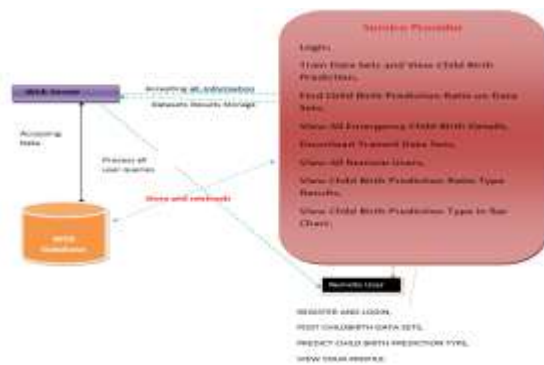


Fig 2 : System Architecture

3.4 Data Flow Diagram : Whenever a new system is developed, user training is required to educate them about the working of the system so that it can be put to efficient use by those for whom the system has been primarily designed. For this purpose the normal working of the project was demonstrated to the prospective users. Its working is easily understandable and since the expected users are people who have good knowledge of computers, the use of this system is very easy.

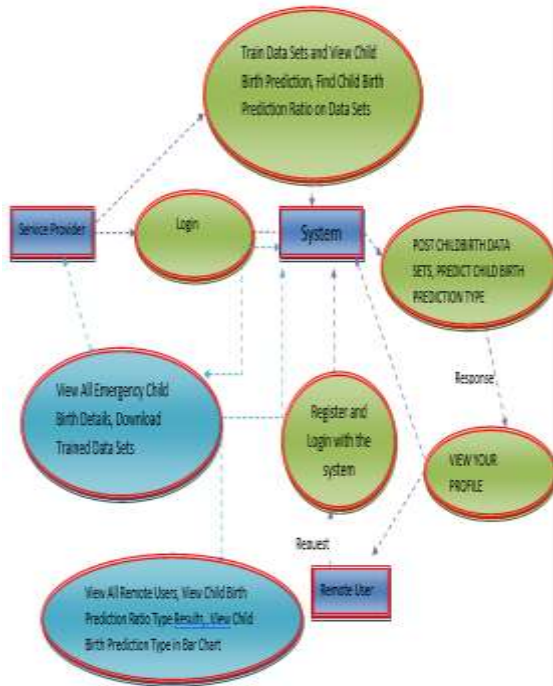


Fig 3: Data Flow Diagram

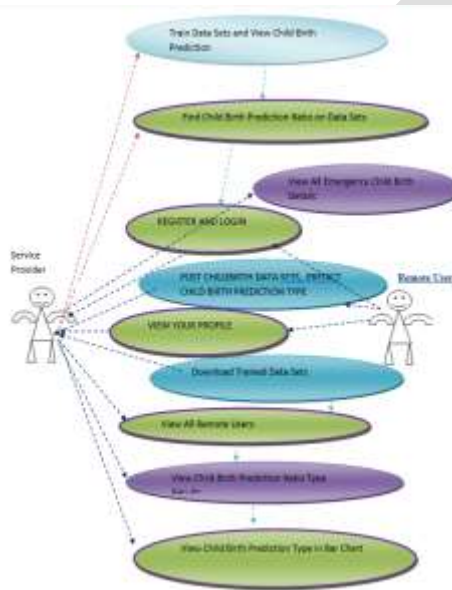


Fig 4 Use Case UML Diagrams

5. CONCLUSION

For the safety of both moms and newborns, selecting the optimal delivery methods is essential. However, research is still needed to determine the optimum aspects to take into account. As a result, this research conducted a thorough empirical investigation, divided all potential features into several groups, and then used a feature selection technique of machine

Mr.K.Vijaya Babu *et. al.*, / International Journal of Engineering & Science Research learning. Later, the best method for predicting the best birthing model with the fewest characteristics was chosen using the results of applying five machine-learning algorithms to combinations of these categories (classes). Performance for several feature classes demonstrated the effectiveness of the methodology used in this investigation.

6. FUTURE SCOPE

The elements that are essential for forecasting delivery modes must be explored in more detail, and it must be determined which supporting factors might boost the precision of such predictions. Similar to this, research should identify whether algorithms or methodologies may give more accuracy when taking into account a certain collection of traits, since it is uncommon to be able to analyse all of a pregnancy's attributes before or during birthing.

7. REFERENCES

- [1] M. Arora. 6 Different Types of Delivery Methods You Must Know. Accessed: Jul. 11, 2020.
- [2] I. Hendler, M. Kirshenbaum, M. Barg, S. Kees, S. Mazaki-Tovi, O. Moran, A. Kalter, and E. Schiff, "Choosing between bad, worse and worst: What is the preferred mode of delivery for failure of the second stage of labor?" *J. Maternal-Fetal Neonatal Med.*, vol. 30, no. 15, pp. 1861–1864, Aug. 2017.
- [3] K. Gregory, S. Jackson, L. Korst, and M. Fridman, "Cesarean versus vaginal delivery: Whose risks? Whose benefits?" *Amer. J. Perinatol.*, vol. 29, no. 1, pp. 07–18, Jan. 2012.
- [4] S. Buckley, *Gentle Birth, Gentle Mothering: A Doctor's Guide to Natural Childbirth and Gentle Early Parenting Choices*. San Francisco, CA, USA: Celestial Arts, 2013.
- [5] I. M. Gaskin, *Ina May's Guide to Childbirth*. New York, NY, USA: Bantam, 2003.
- [6] S. McCutcheon, *Natural Childbirth the Bradley Way: Revised Edition*. Baltimore, MD, USA: Penguin, 2017.
- [7] E. L. Ryding, "Investigation of 33 women who demanded a cesarean section for personal reasons," *Acta Obstetrica et Gynecol. Scandinavica*, vol. 72, no. 4, pp. 280–285, Jan. 1993.
- [8] T. Boerma, C. Ronsmans, D. Y. Melesse, A. J. D. Barros, F. C. Barros, L. Juan, A.-B. Moller, L. Say, A. R. Hosseinpoor, M. Yi, D. de Lyra Rabello Neto, and M. Temmerman, "Global epidemiology of use of and disparities in caesarean sections," *Lancet*, vol. 392, no. 10155, pp. 1341–1348, Oct. 2018.

- [9] V. M. Allen, C. M. O'Connell, R. M. Liston, and T. F. Baskett, "Maternal morbidity associated with cesarean delivery without labor compared with spontaneous onset of labor at term," *Obstetrics Gynecology*, vol. 102, no. 3, pp. 477–482, Sep. 2003.
- [10] S. Liu, R. M. Liston, K. S. Joseph, M. Heaman, R. Sauve, and M. S. Kramer, "Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term," *Cmaj*, vol. 176, no. 4, pp. 455–460, Feb. 2007.
- [11] R. M. Silver, M. B. Landon, D. J. Rouse, K. J. Leveno, C. Y. Spong, E. A. Thom, A. H. Moawad, S. N. Caritis, M. Harper, R. J. Wapner, Y. Sorokin, M. Miodovnik, M. Carpenter, A. M. Peaceman, M. J. O'Sullivan, B. Sibai, O. Langer, J. M. Thorp, S. M. Ramin, and B. M. Mercer, "Maternal morbidity associated with multiple repeat cesarean deliveries," *Obstetrics Gynecol.*, vol. 107, no. 6, pp. 1226–1232, Jun. 2006.
- [12] Mr.K.Vijaya Babu, A Brain-Inspired Cognitive Control Framework for Artificial Intelligence Dynamic System. In: Kumar, A., Mozar, S., Haase, J. (eds) *Advances in Cognitive Science and Communications. ICCCE 2023. Cognitive Science and Technology*. Springer, Singapore. https://doi.org/10.1007/978-981-19-8086-2_70 .
- [13] Dr.C.N.Ravi A Novel Machine Learning Framework for Tracing Covid Contact Details by Using Time Series Locational data & Prediction Techniques.(Volume: 11 Issue: 2s, DOI: <https://doi.org/10.17762/ijritcc.v11i2s.6046>, Article Received: 26 November 2022, Revised: 24 December 2022 Accepted: 04 January 2023)
- [14]. Dr.C.N.Ravi A Review on Various Cloud-Based Electronic Health Record Maintenance System for COVID-19 Patients.(: https://doi.org/10.1007/978-981-19-8086-2_15, April 2023),
- [15]. Dr.C.N.Ravi, Large-scale system identification using self-adaptive penguin search algorithm(DOI: 10.1049/cth2.12479: , Received: 26 November 2022 Revised: 14 March 2023 Accepted: 17 March 2023, *IET Control Theory Appl.* 2023;1–12.),'
- [16]. Dr.C.N.Ravi,, Machine Learning and IoT based Performance Improvement of Energy Efficiency in Smart Buildings.(DOI: [17].1109/ICSCDS56580.2023.10104874, *IEEE Xplore* Part Number: CFP23AZ5-ART; ISBN: 978-1-6654-9199-0, April 2023).
- [18]. Dr.C.N.Ravi ,Classification of Diabetes using Multilayer Perceptron.(DOI: 10.1109/ICDCECE53908.2022.9793085 (22 July 2022))
- [19] Dr.C.N.Ravi ,ANN and RSM based Modeling of Moringa Stenopetala Seed Oil Extraction: Process Optimization and Oil Characterization (Volume: 11 Issue: 7s DOI: <https://doi.org/10.17762/ijritcc.v11i7s>).

[20]E.suresh babu,A Review on Various Cloud-Based Electronic Health Record Maintenance System for COVID-19 Patients

[21]G.sumalatha, Software Defect Prediction Survey Introducing Innovations with Multiple Techniques. In: Kumar, A., Mozar, S., Haase, J. (eds) Advances in Cognitive Science and Communications. ICCCE 2023. Cognitive Science and Technology. Springer, Singapore.
https://doi.org/10.1007/978-981-19-8086-2_75

[22] A.LakshmiBhargav,A Constructive Feature Grouping Approach for Analyzing the Feature Dominance to Predict Cardiovascular Disease |
https://link.springer.com/chapter/10.1007/978-981-19-8086-2_62

[23] D.navaneetha,Detection of Covid-19 from Chest X-Ray Images Using Resnet50, Inception V3 Deep Learning models, JES Journal,VOL 14, ISSUE O5(S)

[25]Y.prathima, Missing Kid Identification Proof Framework Utilizing Deep Learning And Multiclass Svm

[26] Ms.G S Sravanthi, Detection of Covid-19 from Chest X-Ray Images Using Resnet50, Inception V3 Deep Learning models

[27]. Pankaj Kawadkar, Laxmaiah Mettu, Sumalatha G “ Modeling of Information Extraction in Scientific Articles Using Machine Learning Techniques” Published in JOURNAL OF CRITICAL REVIEWS ISSN- 2394-125 VOL 8, ISSUE 04, 2021.