Chapter 10

Book Chapter

Predictive Modeling and Analysis of Fetal Growth using Linear Regression and Random Forest

Rajendra M. Jotawar *1, Swati S. Sangolli †2, and Vachan K R ‡3

[1]Assistant Professor, Department of MCA, Acharya Institute of Technology, Bangalore
[2]Assistant Professor, Department of MCA, Jain College of Engineering, Belagavi
[3]Department of MCA, Acharya Institute of Technology, Bangalore

Abstract

Low fetal birth weight is a critical concern in pregnancy care, significantly affecting neonatal health and contributing to high infant mortality rates globally. Low birth weight is associated with numerous health complications, such as respiratory distress, infections, and long-term developmental challenges. Early diagnosis of fetal growth issues is crucial, as it enables timely medical interventions to prolong the gestation period, allowing more time for fetal development and increasing the likelihood of a healthier birth weight. This project aims to develop a predictive model to estimate fetal birth weight early in pregnancy, categorizing the results as low ($<$ 2.5 kg), normal (2.5–4.5 kg), or abnormal ($>$ 4.5 kg).

Keywords: Artifical Intelligence. Fetal Weight. Headlock Formula. Regressor.

*Email: rajendra2842@acharya.ac.in Corresponding Author
†Email: swati.sangolli@jainbgm.in
‡Email: vachanr.22.mcav@acharya.ac.in

1 Introduction

In recent times, the incidence of newborns with low birth weight has been on the rise, posing significant health challenges. One critical condition contributing to this issue is Intrauterine Growth Restriction (IUGR), a disorder where the fetus is notably smaller compared to others at the same gestational age. This restriction impedes the growth of the baby's body and organs, leading to a myriad of health complications both at birth and later in life. The healthcare sector has seen substantial benefits from the integration of machine learning (ML) approaches. ML, a branch of artificial intelligence (AI), utilizes extensive data to detect patterns and trends that might be overlooked by human analysis. Consequently, ML-based algorithms emerge as powerful tools, aiding healthcare professionals in making informed decisions. The classification of low, norrmal and abnormal weight is crucial as it allows for timely medical interventions that can improve birth outcomes. To achieve this, we employ various machine learning techniques and algorithms, including Linear Regression and Random Forest Regressor. These algorithms analyze a multitude of factors and data points to provide accurate predictions. Through our research, we have determined that the Random Forest Regressor significantly outperforms Linear Regression in terms of prediction accuracy. This finding underscores the potential of advanced ML techniques in enhancing prenatal care and mitigating the risks associated with low birth weight. By leveraging these technologies, healthcare providers can better monitor fetal development and implement necessary interventions, ultimately improving neonatal health and reducing infant mortality rates.

Many models have been created to predict the fetal weight . The Hadlock formula, is used most by many systems (Hadlock et al., 1985). Other notable models include those by Shepard et al.'s (1982), which utilizes different combinations of ultrasound measurements. Studies, like one by Melamed et al.'s (2009), have evaluated these formulas, highlighting their strengths and limitations. Machine learning (ML) and artificial intelligence (AI) have significantly impacted fetal birthweight prediction in recent years. Mennickent et al.'s (2023) demonstrated the efficacy of regression models and neural networks in analyzing complex, non-linear relationships between ultrasound measurements and birthweight. Teles et al.'s (2021) further showed that machine learning algorithms, such as random forests and support vector machines (SVM), can improve predictive accuracy by leveraging large datasets. These studies underscore AI's potential to enhance prediction precision beyond traditional statistical models. Incorporating maternal characteristics such as age, weight, height, gestational age, and health conditions has proven beneficial for accurate birthweight predictions.For example, maternal diabetes and hypertension are known to affect fetal growth, necessitating their consideration in comprehensive prediction models (Lamain – de Ruiter et al., 2017). Despite advancements, several challenges remain in

Editors: S. Pandikumar, Manish Kumar Thakur, Pallavi M O
DOI:10.48001/978-81-966500-2-5-10 | ISBN: 978-81-966500-2-5

fetal birthweight prediction. Ensuring accuracy is paramount, as errors can lead to inappropriate clinical decisions (Girard et al., 2015). Machine learning (ML) and artificial intelligence (AI) continue to revolutionize fetal birthweight prediction. Hang et al.'s (2021) and others have explored the application of ML algorithms to this domain, demonstrating significant improvements in predictive accuracy. Studies like those by Abdi et al.'s (2024) highlight the potential of regression models, neural networks, and other AI techniques in analyzing complex, non-linear relationships between multiple variables, thereby refining the model.

2 Methodologies

- Data Collection
  The first step involves collection of data from different sources. Ultrasound measurements, including Biparietal Diameter (BPD), Head Circumference (HC), Abdominal Circumference (AC), and Femur Length (FL), are gathered using both traditional 2D and advanced 3D/4D ultrasound imaging to improve accuracy and detail. Lifestyle data like nutrition and physical activity are collected via questionnaires and wearable devices. Integration of Electronic Health Records (EHR) ensures a rich source of patient history and prenatal care data, adhering to data privacy and security regulations.
- Data Preprocessing
  Data preprocessing is important for checking the quality of the data which is collected. This step contains cleaning the data to remove errors and inconsistencies, and also handling missing data or missing values. Engineering is done to create new variables by using existing ones. The data is then segmented into training, validation, and test sets, typically in a 70:15:15 ratio, to facilitate model development and evaluation.
- Model Development
  Model development involves creating both traditional and advanced predictive models. Established models, like the Hadlock and Shepard formulas, are used as base. Machine learning models, including regression models like Linear Regression and advanced algorithms like Random Forests and Gradient Boosting Machines, are developed and trained. Deep learning models are designed for very complex and difficult pattern recognition. Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks may be used if temporal data is significant.
- Model Training and Evaluation
  In this phase, models are trained using the training dataset, optimizing hyperparameters through techniques like Grid Search or Random Search. Cross-validation ensures robustness and prevents overfitting. performance is checked using the validating dataset, using metrics like MAE, Root Mean Square Error (RMSE). The best-

DOI:10.48001/978-81-966500-2-5-10 | ISBN: 978-81-966500-2-5

performing model is selected based on these evaluation metrics and its clinical relevance, with considerations for interpretability in clinical settings.

- Integration with EHR and Mobile Technologies
  To enhance the system's functionality, integration with EHR and mobile technologies is essential. Interfaces are developed for easy integration with EHR systems, giving real-time data access and updates. Mobile applications and wearable devices are integrated to collect ongoing maternal health data, such as vital signs and activity levels, which feed into the prediction model to provide continuous monitoring and updates.
  The system goes through various validation and tests to check its effectiveness. Prospective trials are held to check the model's predictive accuracy in real-world settings. Collaborations with healthcare providers enable testing on a diverse patient population. Feedback from these tests is used to refine the user interface and functionality.
- Deployment
  Deployment involves implementing the finalized system in clinical environments, ensuring compatibility with existing workflows. Training is provided for healthcare providers to use the system effectively. Continuous monitoring mechanisms are put in place to maintain model accuracy over time. Regular updates with new data and improvements based on feedback and advancements in the field are essential for the system's sustained effectiveness

3 Results

Our predictive model's results show notable gains in accuracy over conventional ultrasound-based techniques for calculating fetal birth weight. Traditional models such as the Hadlock and Shepard formulas were routinely surpassed by the machine learning algorithms, especially the Random Forest Regressor. Through the integration of an extensive dataset comprising maternal health data, such as exercise levels, nutrition, and medical history, together with ultrasound readings, our algorithm was able to capture the intricate patterns of fetal growth that are frequently overlooked by traditional methods. Its resilience in managing nonlinear relationships between the features was highlighted by the Random Forest Regressor's lower Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) during validation when compared to Linear Regression.By giving more precise anatomical data, the use of modern ultrasound imaging technologies, such as 3D/4D imaging, substantially improved the accuracy of our predictions. This reduced the possibility of misclassification and made it possible for more precisely focused medical interventions by classifying fetal birth weight into the clinically relevant groups of low, normal, and abnormal birth weights.

DOI:10.48001/978-81-966500-2-5-10 | ISBN: 978-81-966500-2-5

When the prediction model was tested on a wide range of patients, it produced encouraging findings. We carried out prospective trials in partnership with healthcare professionals to evaluate the efficacy of the model in diverse clinical contexts. The model was especially helpful in the early stages of pregnancy, when standard approaches frequently fail to identify possible growth concerns, according to feedback from these trials. Real-time information on fetal health were made possible by the continuous monitoring capabilities made possible by the combination of mobile technology with EHR systems. This gave healthcare clinicians a flexible tool to modify prenatal care plans as necessary. Because of the model's high accuracy in predicting birth weights, clinical decision-making was enhanced, and the likelihood of unfavorable newborn outcomes—like respiratory distress—associated with low birth weight was decreased.

4 Conclusion

There are difficulties in correctly diagnosing fetal health since current ultrasound-based methods for monitoring fetal growth frequently misclassify about 15% of fetuses as short for gestational age (SGA). It is challenging to detect growth restriction early on using conventional models, which is typically caused by an insufficient supply of nutrients and oxygen. A accurate diagnostic technique is still elusive despite the several ultrasonography formulas that have been offered; early detection is crucial to better outcomes. Our project seeks to improve the prediction of fetal birth weight by utilizing machine learning (ML) techniques like Random Forest Regressor and Linear Regression. Our approach provides more accurate predictions by combining these algorithms with extensive maternal health data, categorizing birth weights into low, normal, and abnormal categories.This approach highlights the transformational potential of machine learning (ML) in fetal health evaluations by helping healthcare providers make better decisions, improve early detection, and deliver better prenatal treatment.

References

Abdi, E., Ali, M., Santos, C. A. G., Olusola, A., & Ghorbani, M. A. (2024). Enhancing groundwater level prediction accuracy using interpolation techniques in deep learning models. Groundwater for Sustainable Development, 26. https://doi.org/10.1016/j.gsd.2024.101213

Girard, M. J., Dupps, W. J., Baskaran, M., Scarcelli, G., Yun, S. H., Quigley, H. A., Sigal, I. A., & Strouthidis, N. G. (2015). Translating ocular biomechanics into clinical practice: Current state and future prospects. Current Eye Research, 40(1), 1–18. https://doi.org/10.3109/02713683.2014.914543

DOI:10.48001/978-81-966500-2-5-10 | ISBN: 978-81-966500-2-5

Hadlock, F. P., Harrist, R. B., Sharman, R. S., Deter, R. L., & Park, S. K. (1985). Estimation of fetal weight with the use of head, body, and femur measurements-A prospective study. American Journal of Obstetrics and Gynecology, 151(3), 333–337. https://doi.org/10.1016/0002-9378(85)90298-4

Hang, Y., Liu, T., Chen, L., & Wang, Q. (2021). Prediction of fetal weight using machine learning models with sonographic measurements. Journal of Medical Imaging and Health Informatics, 11(3), 829–835.

Lamain – de Ruiter, M., Kwee, A., Naaktgeboren, C. A., Franx, A., Moons, K. G. M., & Koster, M. P. H. (2017). Prediction models for the risk of gestational diabetes: a systematic review. Diagnostic and Prognostic Research, 1(1). https://doi.org/10.1186/s41512-016-0005-7

Melamed, N., Yogev, Y., Meizner, I., Mashiach, R., Bardin, R., & Ben-Haroush, A. (2009). Sonographic Fetal Weight Estimation. Journal of Ultrasound in Medicine, 28(5), 617–629. https://doi.org/10.7863/jum.2009.28.5.617

Mennickent, D., Rodríguez, A., Opazo, M. C., Riedel, C. A., Castro, E., Eriz-Salinas, A., Appel-Rubio, J., Aguayo, C., Damiano, A. E., Guzmán-Gutiérrez, E., & Araya, J. (2023). Machine learning applied in maternal and fetal health: a narrative review focused on pregnancy diseases and complications. Frontiers in Endocrinology, 14. https://doi.org/10.3389/fendo.2023.1130139

Shepard, M. J., Richards, V. A., Berkowitz, R. L., Warsof, S. L., & Hobbins, J. C. (1982). An evaluation of two equations for predicting fetal weight by ultrasound. American Journal of Obstetrics and Gynecology, 142(1), 47–54. https://doi.org/10.1016/S0002-9378(16)32283-9

Teles, G., Rodrigues, J. J., Rabêlo, R. A., & Kozlov, S. A. (2021). Comparative study of support vector machines and random forests machine learning algorithms on credit operation. Software - Practice and Experience, 51(12), 2492–2500. https://doi.org/10.1002/spe.2842

DOI:10.48001/978-81-966500-2-5-10 | ISBN: 978-81-966500-2-5