Similarity Report

Original Document:

Smart Fetal Growth Monitoring, Real-Time Weight Prediction in High-Risk Pregnancies: A Systematic Literature Review

[1] Anusha K H,
[1] Bhuvan M,
[1] Deepika R S,
[1] Deekshith Chandra
Information and Science and Engineering, Malnad College of Engineering
Hassan-573202, India
Email id: anushaharish48@gmail.com, bhuvibhuvan1624@gmail.com, deepikars168@gmail.com, deekshithchandra09@gmail.com

Abstract— The newborn's low birth weight is one of the most important issues in prenatal care since it can negatively impact the infant's health and, in more severe cases, even result in its death. This condition is the reason behind the high infant mortality rates seen globally. Methods of artificial intelligence, especially those based on machine learning (ML), can predict health problems that might occur at birth as well as during the entire gestation period. Therefore, our study proposes to analyze a number of machine learning (ML) techniques that can be used to predict if a fetus would be born weighing less than what is expected for its gestational age. The importance of identifying fetal development problems early on is emphasized by the possibility of extending gestation days through timely intervention. Using such an intervention, A decrease in infant morbidity and death would result from the potential to raise fetal weight at birth. Therefore, in this research, we will forecast the fetal birth weight at an early stage and classify them as low weight if their weight is less than 2.5 kg, normal weight if their weight is more than 2.5 kg but less than 4.5 kg, and abnormal weight if their weight is more than 4.5 kg.We used machine learning techniques and algorithms, such as Linear and Random Forest Regressor, to estimate the fetal birth weight in this case; Random Forest Regressor predicted the weight more precisely than Linear Regression.

Keywords: random forest regression, linear regression, neonatal morbidity, gestational age, and infant death rates.

I. INTRODUCTION

Intrauterine growth restriction (IUGR) and low birth weight (LBW) are major issues in prenatal care that have a big impact on the health of the newborn. Low oxygen levels, poor Apgar scores, respiratory issues from meconium aspiration, and hypoglycemia are some of the consequences that can result from IUGR, in which the fetus is smaller than others of the same gestational age. In extreme situations, IUGR may cause long-term growth issues or even fetal death. The most severe prenatal development problems are known to be caused by persistent arterial hypertension, and maternal hypertension is the primary cause of IUGR. Increased chances of newborn death, stunted growth, mental retardation, learning disabilities, and chronic illnesses like obesity, diabetes, and heart disease have all been associated with low birth weight. Similarly, challenges including cesarean delivery, extended labor, hemorrhage, and trauma during delivery, as well as the risk of infant hypoxia and death, are linked to large birth weight (macrosomia).

In this regard, fetal birth weight is predicted and analyzed using machine learning (ML) approaches, including Linear Regression and Random Forest Regressor, which offer insightful information that can assist healthcare professionals in making defensible decisions. Machine learning (ML) can help predict low birth weight and identify issues early by seeing trends in huge datasets. This enables prompt interventions and individualized treatment plans. Predicting the fetal birth weight early in pregnancy and categorizing it into three groups—low birth weight (less than 2.5 kg), normal birth weight (2.5–4.5 kg), and abnormal birth weight (more than 4.5 kg)—is the goal of this effort. Reducing maternal and neonatal morbidity and death requires early diagnosis of these disorders.

II. RELATED WORKS

A. Paper Title: Prediction of Weight Range of Neonate Using Machine Learning Approach, Authors: Aleem Adeeba, Banujan Kuhaneswaran, and B.T.G.S. Kumara, Year of publication: 2022

Description: The study predicts newborn birth weight classes using machine learning with data from 500 Sri Lankan women. Five algorithms—ANN, Naive Bayes, SVM, Logistic Regression, and Decision Trees—are used to find cost-effective solutions for newborn health prediction in low-resource settings.

Methodology: Researchers identified 16 factors affecting neonatal birth weight. Python preprocessed data from 500 prenatal records in Sri Lanka, splitting it into 80% training and 20% testing. Accuracy, precision, recall, and F1 scores were evaluated for five models: ANN, Naive Bayes, SVM, Logistic Regression, and Decision Trees. ANN used the Adam optimizer and binary cross-entropy loss.

Limitations: Generalizability and comprehension of neonatal health outcomes may be impacted by the study's limitations, which include its small dataset (500 samples), regional focus on Sri Lanka, and the simple classification of infant weights (<3,200 g and "e3,200 g).

Key Insights: The study demonstrated the use of machine learning models for predicting neonatal birth weight, focusing on low-resource settings. Data preprocessing and model evaluation with metrics like accuracy and F1 scores highlighted the effectiveness of these models. However, the small dataset, regional focus, and simplified classification limited generalizability.

B. Paper Title: Machine learning-based approach for predicting low birth weight, Authors: Ranjbar, A., Montazeri, F., Mehrnoush, V., Farashah, M. V., Darsareh, F., & Roozbeh, N, Year of publication : 2023

Description: Using data from 8,853 births in Iran's IMaN Net, Ranjbar et al. (2023) investigated machine learning techniques for forecasting low birth weight (LBW), a significant newborn risk factor. They assessed eight machine learning algorithms, such as deep learning, random forest, and XGBoost. In 14.5% of instances, LBW (<2,500 g) was present. The study emphasizes how ML can help improve prenatal care via early detection of high-risk pregnancies.

Methodology: The study examined IMaN Net data on singleton pregnancies that lasted longer than 24 weeks from 2020 to 2022. Using 10-fold cross-validation, AUROC, accuracy, and F1 score were used to assess eight machine learning models, including XGBoost and SVM. Important indicators such as LBW history and gestational age were found.

Limitations: The study's shortcomings include excluding multiple pregnancies and anomalies, relying solely on one database, missing important variables like maternal BMI, and XGBoost's AUROC values suggesting room for improvement in accuracy.

Key Insights: XGBoost obtained an AUROC of 0.79, 79% accuracy, and 87% precision. Although adding variables could increase accuracy, key indicators like as gestational age and LBW history highlight ML's potential to improve LBW forecasts.

C. Paper title:A Systematic review on applications of machine learning for fetal birth weight prediction, Authors: Sasidhar Babu, S., and Keerthana, P., Year of publication : 2024

Description: Keerthana and Sasidhar Babu's (2024) review of 85 studies highlights ML's superior accuracy in fetal weight prediction, with techniques like XGBoost, random forest, and SVM achieving up to 100% accuracy using maternal and fetal data.

Methodology: A systematic evaluation of peer-reviewed publications from Elsevier and Springer was part of the process, with an emphasis on research employing machine learning techniques such as CNNs, SVMs, and random forests. Non-ML techniques were removed, and only recent English-language papers were included.

Limitations: Dataset heterogeneity, the challenge of extrapolating results, and the absence of features such as maternal BMI in certain models are among the limitations. Reproducibility is also impacted by variations in study technique and data quality.

Key Insights: : Important discoveries highlight machine learning's contribution to predicting prenatal health, increasing precision, facilitating individualized treatment, and lowering newborn hazards. The accuracy of neural networks was the highest, demonstrating the scalability and

dependability of machine learning.
D. Paper title: Machine learning improves early prediction of small for gestational age births and reveals nuchal fold thickness as unexpected predictor, Authors: Biswas, A., Saw, S. N., Lee, H. K., Mattar, C. N. Z., & Yap, C. H., Year of publication : 2021
Description: Using ultrasound data from the second trimester, the study investigates machine learning models such as random forest, SVM, and MLP to predict small-for-gestational-age (SGA) babies. Nuchal fold thickness was found to be a significant predictor in 347 pregnancies when ML performed better than clinical recommendations.
Methodology: Using preprocessed data for homogeneity, the methodology trained three machine learning models on 16 input factors, such as maternal demographics, fetal biometry, and Doppler indices. Cross-validation was used to assess the models and compare them to clinical diagnosis. SVM outperformed clinical diagnosis by 25% and had the highest accuracy (83%) in predicting severe SGA.
Limitations: The small sample size, dependence on single-institution data, and exclusion of maternal history or biomarkers are among the limitations. Data overlaps and intra-observer variability are issues.
Key Insights: With NF thickness being essential for SGA prediction, key findings demonstrate ML's superior performance over clinical guidelines and underscore ML's capacity to identify new predictors in fetal growth monitoring.
E. Paper title: Birthweight Range Prediction and Classification: A Machine Learning-Based Sustainable Approach, Authors: S.Y. Ajibi, R.B. Alotaibi, D.A. Alabbad, et al., Year of publication : 2024
Description: The study addresses the prediction and classification of fetal birth weight ranges using machine learning (ML) to mitigate health risks for mothers and newborns. It incorporates clinical data from Saudi Arabia and publicly available datasets to create models that classify weights as low, normal, or high, demonstrating improved accuracy compared to traditional ultrasound methods.
Methodology: In an effort to lower health risks, the study predicts and categorizes fetal birth weight ranges using machine learning (ML). It surpasses conventional ultrasound techniques by combining Saudi Arabian clinical data with publicly accessible datasets to generate models that categorize weights as low, normal, or high.
Limitations: The study's generalizability is restricted by its dependence on particular datasets. Even with excellent accuracy, the models might require additional fine-tuning for practical application. Key influences on birth weight may be missed if more general demographic or genetic factors are excluded.
Key Insights: On the Saudi dataset, Random Forest's accuracy was 96%, while Extra Trees' was 98%. Weight and blood pressure were important indicators for both the mother and the newborn, demonstrating the promise of machine learning in prenatal treatment.

F. Paper Title: A Comparative Study of Machine Learning Algorithms for Predicting Weight Range of Neonate, Authors: Verma, R., and K. Sharma, Year of publication:2022

Description: In order to lower delivery risks and improve early diagnosis of abnormal birth weights, this research examines eight machine learning algorithms that predict neonatal weight ranges using maternal health markers.

Methodology: The study tested algorithms like SVM, Random Forest, and Gradient Boosting using maternal health data, such as blood pressure, age, gestational age, and BMI. Missing values and normalization were managed via data preparation, and robust comparison was ensured by assessment metrics like accuracy, precision, and recall.

Limitations: Generalizability was impacted by the dataset's small size and lack of diversity. Predictive ability was diminished by missing variables such as genetic and lifestyle factors, unbalanced data, and a lack of outside validation.

Key Insights: Simpler models performed worse than Random Forest and Gradient Boosting. Maternal BMI and gestational age were important indicators; ensemble approaches improved consistency and highlighted the need for more comprehensive information.

III. METHODOLOGY

Attribute selection: In attribute selection, clinically significant factors are identified through expert interviews, validated and expanded upon through a comprehensive literature review, and the parameters that have the greatest impact on fetal weight—such as maternal health, ultrasound metrics, and prenatal care indicators—are finalized using statistical or machine learning techniques.

Data Collection: A number of variables and past data must be acquired throughout the data collecting process in order to predict fetal birth weight with any degree of accuracy. Maternal health data, including age, height, weight, blood pressure, blood sugar, and hemoglobin levels, are collected to establish a baseline for understanding the mother's physical condition and its potential impact on the fetus's growth. Fetal development data, such as fundal height, fetal movements, and fetal gender, are also provided to give information on the fetus's growth trajectory and overall health. The mother's medical history is examined, including records of hypertension, diabetes, and prior pregnancies, to identify patterns or risk factors that might result in anomalies in fetal growth.

Data Preprocessing: Data Cleaning: Use methods like mean/mode imputation or interpolation to fill in missing values. Find and eliminate outliers and duplicate records that can distort the findings.Data Transformation: To guarantee consistency, normalize or scale numerical variables using a standard scaler or min-max scaling. To make categorical variables machine-readable, use label encoding or one-hot encoding. Feature Engineering: To find and choose the most important features for prediction, apply methods like principal component analysis (PCA), correlation analysis, or recursive feature elimination (RFE).

Data Splitting: Assign 20% of the dataset for model assessment and 80% of the dataset for model training. Make sure that the training and testing subsets preserve the same class distribution (low, normal, and abnormal birth weights).

Model Selection: Based on research findings, model selection entails assessing many machine learning algorithms, including Random Forest, SVM, Logistic Regression, and Deep Learning. Examine their appropriateness for the dataset, accuracy, and performance. Select the optimal algorithm based on prediction accuracy and project needs.

Model Training and Evaluation: Model training entails employing GridSearchCV or RandomizedSearchCV to optimize the performance of certain algorithms that have been trained on the dataset. Analyze models using test data, evaluating how well they predict birth weights and categorize data using measures such as MAE and MSE to guarantee accurate predictions.

Deployment: Implement the model by developing an intuitive online or mobile application that allows users to enter data, see forecasts, and get useful insights.

Fig. 1. Flow Chart

IV. ALGORITHMS USED

A. Support Vector Machine (SVM)

SVM is a robust supervised learning algorithm primarily used for classification tasks, including both binary and multi-class problems. It operates by identifying the optimal hyperplane that best separates data points into distinct categories. SVM is particularly effective in high-dimensional spaces and when the dataset exhibits a clear margin of separation. By employing kernel functions such as linear, polynomial, and radial basis function (RBF), SVM can also handle non-linear relationships in the data. Its excellent generalization ability makes it a popular choice for applications like image recognition and prediction.

Fig. 2. SVM Algorithm

B. K-Nearest Neighbors (KNN).

KNN is an instance-based learning algorithm that classifies data points based on the proximity of other data points. The algorithm identifies the 'k' closest neighbors to a given data point and assigns the most frequent class among them. KNN is easy to implement and does not require a formal training phase, making it ideal for small datasets. However, its performance is highly dependent on the choice of 'k' and the distance metric (e.g., Euclidean or Manhattan distance). KNN performs well when data points are grouped into distinct clusters and is commonly used in applications like prediction and recognition .

Fig. 3. KNN Algorithm

C. XGBoost

XGBoost (Extreme Gradient Boosting) is a powerful machine learning algorithm that excels in predictive accuracy and efficiency. It is an ensemble method that builds a series of decision trees using gradient boosting to minimize errors. XGBoost is particularly effective for job prediction tasks, such as forecasting job placement success or identifying trends in industry demands. It works well with structured datasets, handling missing values and noisy data effectively. XGBoost's speed and performance make it a top choice for predicting career outcomes and identifying key factors influencing job success.

Fig. 4. XGBoost Algorithm

D. Decision Tree

Random forest is a supervised learning algorithm. It can be used both for classification and regression. It is also the most flexible and easy to use algorithm. A forest is comprised of trees. It is said that the more trees it has, the more robust a forest is. Random forests creates decision trees on randomly selected data samples, gets prediction from each tree and selects the best solution by means of voting. It also provides a pretty good indicator of the feature importance. It technically is an ensemble method (based on the divide-and-conquer approach) of decision trees generated on a randomly split dataset. This collection of decision tree classifiers is also known as the forest.

E. ANN

Artificial Neural Networks (ANNs) are computer models made of linked neurons arranged in layers that are modeled after the structure of the human brain. To generate predictions, they use activation functions, weights, and biases to process inputs. ANNs use gradient descent and backpropagation to minimize errors. They are widely used in tasks like image recognition, natural language processing, and forecasting.

V. PROPOSED SYSTEM

In high-risk pregnancies, the suggested approach seeks to forecast the fetal birth weight and categorize it into three groups: low, normal, and abnormal. In order to provide accurate and dependable predictions, the system makes use of sophisticated machine learning techniques, particularly Linear Regression and Random Forest Regressor. Both algorithms will be employed for prediction, and the accuracy attained by each approach will be compared.

Key Features of the System:

Prediction:Fetal weight is categorized as:

Low Birth Weight (LBW): Less than 2.5 kg.

Normal Birth Weight (NBW): Between 2.5 kg and 4.5 kg.

Abnormal Birth Weight (ABW): Greater than 4.5 kg.

Data Utilization:
Makes use of a dataset including fetal statistics, including ultrasound measurements, maternal health status, and gestational age. manages preprocessing duties to guarantee data quality, such as eliminating duplicate or missing items.
End-to-End Functionality:
Frontend: Developed with HTML and Flask, this area allows users to enter parameters.
Backend: Uses a trained machine learning model to process the inputs and provides real-time predictions.
Model Lifecycle:
Training and testing sets of data are separated. Using the designated ML algorithms, prediction models are created using the training data. Before being deployed, the model is tested to determine its correctness.
System Workflow

Data Collection and Preprocessing: To increase prediction accuracy, the dataset is loaded and cleaned. Missing values are efficiently handled, and irrelevant data is filtered.
Model Training: The Random Forest Regressor and Linear Regression are the two machine learning techniques used. The model is trained using pre-processed data, which guarantees a high degree of generalization.
Model Evaluation: Accuracy for each ML method is computed. According to the studies have been made highest accuracy must be achieved by The Random Forest.
User Interaction: Using an intuitive interface, the system gathers user-inputted parameters and generates predictions in real-time, presenting the category and the estimated fetal birth weight.
Output: The output shows the predicted fetal birth weight and categorizes the result into LBW, NBW, or ABW categories.

Fig. 5. System Architecture

VI. INTERMEDIATE CONCLUSION

Because of its strong design and efficient use of machine learning methods, including Linear Regression and Random Forest Regressor, the suggested system for fetal birth weight prediction has been chosen for deployment. To guarantee precise forecasts and real-time outcomes, the system's architecture combines an intuitive user interface with a robust backend.
Key highlights include:
Algorithms: Random Forest Regressor is the primary choice due to its superior accuracy.
System Design: Modular and scalable, combining Flask-based frontend and backend ML models.
Impact: Enables early detection of risks, improving maternal and neonatal health outcomes.
We are now set to proceed with system development, focusing on data preprocessing, model training, and system integration. This marks a significant step toward delivering a functional and impactful solution.

REFERENCES

[1] A. Motwani et al., “Soil Analysis and Crop Recommendation using Machine Learning,” 2022.
[2] P. Pudumalar et al., “Crop Recommendation System for Precision Agriculture,” 2016.
[3] M. R. Madhumathi et al., “Prediction of Soil Fertility using ML Algorithms and Fertilizer Recommendation System,” 2023.
[4] P. Sharma et al., “Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning,” 2023.
[5] W. Albattah et al., “A Novel Deep Learning Method for Detection and Classification of Plant Diseases,” 2021.
[6] S. Babu, “A software model for precision agriculture for small and marginal farmers,” in Proc. 2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite, pp. 352-355, doi: 10.1109/GHTCSAS.2013.6629944.
[7] S. Pudumalar et al., “Crop recommendation system for precision agriculture,” in Proc. 2016 Eighth Int. Conf. Adv. Comput, pp. 32-36, doi:10.1109/ICoAC.2017.7951740
[8] R. K. Rajak et al.,“Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique,” Int. Res. J. of Eng. and Technol., issue 12, vol. 4, pp. 950-953, Dec 2017
[9] A.K.M. Reddy, S. Chithra, H.M. Hemashree and Thanu K., “Soil Classification and Crop Suggestion using Machine Learning,” Int. J. for Res. in Appl. Sci. and Eng. Technol., issue 7, vol. 8, pp. 1625-1628, Jul 2020
[10] M. S. Rao, A. Singh, N. V. S. Reddy, and D. U. Acharya, ‘‘Crop prediction using machine learning,’’ J. Phys., Conf. Ser., vol. 2161, no. 1, 2022, Art. no. 012033

Compared Document:

2022 3 rd Int. Conf. on Innovations in Science, Engineering and Technology (ICISET) 26-27 February 2022, Chittagong, Bangladesh 978-1-6654-8397-1/22/$31.00 ©2022 IEEE *Prediction* *of* *Weight* *Range* *of* *Neonate* *Using* Machine Learning Approach Saleem Adeeba Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka Belihuloya, Sri Lanka sadeeba@std.appsc.sab.ac.lk Banujan Kuhaneswaran Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka Belihuloya, Sri Lanka bhakuha@appsc.sab.ac.lk B.T.G.S Kumara Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka Belihuloya, Sri Lanka btgsk2000@gmail.com Abstract— The weight of a neonate at birth is closely linked to mortality risk in the first year and developmental issues in infancy, and the chance of developing numerous diseases in adulthood. In epidemiological studies, *birth* weight is frequently considered the causal pathway to these health outcomes. Not all people have the facility to scan their fetus’s weight before birth. In our research, we proposed a method to predict the weight class of the newborn baby by using existing data of pregnant women for the first six months. We *identified* 16 factors after an extreme literature review and expert opinion. Gradivity, age of pregnant woman, blood group, pregnant woman’s employment, history of abortion, consanguinity, etc., *are* *some* of the factors identified. A dataset of 500 pregnant women was collected and preprocessed for applying the machine learning models. Once the preprocessing was done, the dataset was processed with Machine Learning (ML) models. After completion of data preparation, ML algorithms were used for *the* *training* and testing data set. This research utilized five ML algorithms to predict newborn baby weight at the six months of pregnancy. ANN and NB models have given accuracy above 70%, SVM and logistic regression algorithms provide accuracy above 60%, and Decision Tree provides accuracy below 60%. Keywords—machine learning, newborn, neonate weight, ANN I. INTRODUCTION The shift from fetal to the neonate (also known as newborn), which occurs at delivery, is complicated and must occur swiftly to survive on its own [1]. The fetus prepares for the transition by creating hormones (cortisol, adrenaline, and thyroid hormones) that will spike at delivery, allowing the neonate to resume regular blood sugar, body temperature swiftly, *and* blood pressure levels [2, 3]. Baby *birth* weight is an important measurement to protect maternal and infant safety [4]. Agreeing to the World Health Organization (WHO), neonate weights between 3,000 g and 3,500 g are considered as average healthy baby birth weight, and baby weight of less than 2,500 g is assessed as *Low* Birth Weight (LBW), regardless of how far along the pregnancy was [5]. It signals a neonate’s chances of being alive, growth, long-term health, and psychological development. Therefore, it’s crucial to identify the factors that affect the baby's birth weight before pregnancy and until birth [6]. For example, *based* on the previously studied research, the factors that affect baby birth weight are commonly identified as genetics, parents' age, weight gains during pregnancy, length of gestation, gender of baby, number of babies, and so on [7, 8]. Further, according to the WHO recommendations, pregnancy weight gain with a good nutritional status is 10,000 g to 14,000 g (10 kg-14 kg) [9]. Tracking a baby's weight from birth is essential to ensuring healthy and growing normally [10]. Smaller babies have a higher risk of birth complications and a higher risk of having illnesses and other medical problems later in life [11]. LBW babies are also more probability of dying within the first twenty-eight days of their lives. That's why every country has a baby scale on maternity wards (and why health visitors always have them); a baby is weighed almost immediately after birth. Baby scales must be highly accurate to ensure that they deliver a dependable weight reading with pinpoint accuracy [12]. Predicting the baby's weight at the early stage undoubtedly determines the childbirth method appropriately [13]. It will be vital to identify *risk* factors that influence baby birth weight, such as (i) LBW, (ii) the higher *the* *risk* of infant mortality, (iii) labour complication, (iv) neonate survival, and (v) multiple birth complication [14]. Further, it will inform a baby’s health in advance to the pregnant woman and medical practitioners, which will allow them to evaluate the nutrition level of the pregnant woman. More sophisticated techniques/equipment provide a way to predict the baby's birth weight [15, 16] in an appropriate, accurate, and standard manner. But these sophisticated equipment’s are only used in developed countries. Developing countries like Sri Lanka lack enough cost, technology, resources, and trained workforce to use this equipment and are restricted to predicting the birth weight of every child [17]. A country like Sri Lanka manually maintains the pregnancy records details to inform the baby’s health during pregnancy [18]. However, in developed countries, they use automated technology to predict the health and weight of the baby in advance that gives an accurate calculation [17]. Considering the issues and importance of the above study, there is a need for an alternative prediction technique to find the birth weight of every child in every country. Therefore, seeking a substitute method *is* easy to apply, with minimum cost and enough resources. However, the baby *birth* weight prediction cannot be made directly, and it follows rough estimation with some association factors [19] before pregnancy, the pregnancy, until birth. It includes a pregnant woman's weight, height, age, haemoglobin level, sugar level during pregnancy, etc. [20]. This research paper identifies an alternative prediction method of baby birth weight class for every child with the minimum cost at the pregnancy stage at six months. Thus, a baby *weighing* less than 3,200 g is considered one class, and a baby weighing greater than or equal to 3,200 g is regarded as another class. 427 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET) | 978-1-6654-8397-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICISET54810.2022.9775840 Machine Learning (ML) is a study of computer algorithms that can learn and develop on their own with experience and data. It’s often seen as part of artificial intelligence. ML models apply sample data *to* *create* models that can make forecasts without being definitively programmed. Globally, it’s used in broad fields. Such as health, education, agriculture, Information technology, and so on [21]. The main objective of this paper is to utilize ML technologies to predict the baby birth weight class with minimum errors at the end of 6 months. This research paper proposes a baby’s *birth* weight prediction model based on ML algorithms. The model trained with a different 16 factors (attributes) that affect the baby's birth weight. Such as pregnant woman’s BMI, pregnant woman’s age, and haemoglobin levels *are* some of them. II. LITERATURE REVIEW AND RELATED WORKS A baby's *birth* weight is the first thing practitioners consider while the pregnant woman delivers the baby. However, all baby’s *birth* weight is not predefined; it will vary for each neonate. Knowing a baby's weight is a clear *indicator* of the baby’s health. Therefore, knowing the factors affecting a baby’s birth weight consideration is a must. The UK MARSDEN article discusses that the factors affecting baby weights are (i) genetics (ii) age of pregnant woman (iii) number of children (iv) length of pregnancy (v) pregnant woman’s birth baby weight (vi) diet during gestation (vi) pregnant woman’s practice (viii) gender of the baby (ix) parent’s medical condition and (x) ethnicity [7]. Moreover, the Healthgrades article discusses affecting factors *such* *as* blood pressure, diabetes, heart diseases, asthma, kidney disease, lupus, anaemia, and dental health [12]. Also, previous studies were carried out in different countries regarding the components impacting the neonate’s weight. According to Chandra S. Metgud, Vijaya A. Naik, and Maheshwar D. Mallapur finds, factors impacting baby weight include illiteracy among pregnant women, passive smoking exposure, childbearing later in life, shorter intervals between pregnancies, pre low weight kids, the weight of pregnant women, growth of weight, pregnancy-induced hypertension, pregnancy with a significant risk of complications, and antenatal registration beyond the due date. It was carried out in India, Rural Karnataka [8]. One of the publications had been carried out by students in London discussed the most influencing factors for affecting birth baby weight was mentioned as smoking, maternal height and parity, gestation, and the baby's sex [22]. Further, Talie, et al. [23] discusses that infant sex, anaemia during pregnancy, history of LBW, pregnant woman’s employment, and so on are the factors affecting the birth baby weight, it was carried out in Debre Markos Referral Hospital, Northwest Ethiopia [5, 23]. As the literature results mentioned above indicate, a baby’s *birth* weight is influenced by several factors. Some factors may vary from one country to another. Further, the above literature primarily focuses on the reasons for the LBW. As a common understanding, similar influencing factors in each country are mentioned as pregnancy weight, height, age, gender of the baby, and maternal weight. Based on the previous studies and concerns from medical practitioners, this research paper identifies the following 16 factors that affect the baby weight *used* to predict the baby weight at the six months, such as gravidity, age of pregnant woman, blood group of a pregnant woman, pregnant woman’s employment, history of abortion, consanguinity, pre- pregnancy scanning done, history of subfertility, height, weight, and BMI of pregnant woman, haemoglobin level, blood sugar, tracking of the weight of pregnant woman until first six months pregnancy, fundal height, and gender of the baby. ML is an essential technology for the healthcare sector. Thus, it is being used in many aspects of health care, including the invention of emerging health treatments, patient data and information management, and chronic illness therapy. *According* to the New Yorker, computer scientist Sebastian Thrum: “Machines will make the human brain 1,000 times more powerful, just as they made human muscles 1,000 times stronger” [21, 24]. For example, illnesses and therapies are predicted, health risks are forecasted for distinct demographics of people, assist in the management of medical records and workflow, distinguishes between malignancies and normal anatomy, drug development is aided and expenditures are reduced, identify potential clinical trial sites, identifies healthcare gaps and assists pathologists in making more accurate and timely diagnoses, etc. These are some examples *that* can be achieved using ML algorithms [21, 24]. Above mentioned studies have given the importance of ML and ML in healthcare. Therefore, this research paper identifies the ML algorithms to predict *birth* weight at six months of pregnancy. Karthiga, et al. [25] used TensorFlow ML algorithms to predict the newborn's birth weight. It’s emphasized on creating an online application that forecasts newborn weight *based* on the gender of the baby, the number of siblings, length of pregnancy, and age. Tao, et al. [26] focused on *fetal* *birth* weight prediction with measured data using temporal ML. By merging multiple e-medical data with the B-ultrasonic monitoring of pregnant women, this research aims to develop a hybrid *birth* weight prediction classifier using long short- term memory (LSTM) networks. Hussain and Borah [27] had predicted the baby weight *based* on the different types of features of the pregnant woman. In this study, the newborn baby's *birth* weight prediction was carried out using two ML techniques said Gaussian NB and Random Forest. As a result of the literature mentioned above, it has the drawback of the prediction period of the baby weight. Thus, the prediction period of the baby's weight is not mentioned. But this research paper fills the gap by predicting the baby weight of the six months of the pregnancy. Predicting *birth* weight at six months before the baby is born is the most effective strategy to ensure that the neonate receives special attention as soon as feasible. It assists us in making arrangements for doctors and specific facilities before the birth of the kid. Table I shows the comparison of the existing studies in the context of number of attributes, technologies (algorithms) used for prediction and prediction month of the birth weight. 428 TABLE I. COMPARISON OF THE EXISTING STUDIES Paper # of attributes Some attributes names Technologies used Prediction month of the birth weight [25] 7 Weight in pounds, Development weeks, majority etc. ML (TensorFlow) library Not mentioned or after completion of gestational weeks [26] 6 Height, Age, Abdominal Circumference of Pregnant Women etc. LSTM Not mentioned or after completion of gestational weeks [27] 18 Age, Height, Blood Group of mothers etc. Gaussian NB *and* Random Forest Not mentioned or after completion of gestational weeks [28] 9 Smoking, Race, Maternal-Age, Weight etc. ANN Not mentioned or after completion of gestational weeks This research 16 Gravidity, Age, Blood group, Height etc. SVM, ANN, NB, Decision Tree, and Logistic Regression At six months of period III. PROPOSED APPROACH Fig 1 shows the methodology graph for baby weight prediction using ML algorithms at six months of pregnancy. A. Attributes selection: Several interviews were conducted with experts, and along with the help of experts’ comments and from the extreme literature survey, 16 factors were identified that affect the neonate weight. Thus, gravidity, age of pregnant woman, blood group, pregnant woman’s employment, history of abortion, consanguinity, pre-pregnancy scanning, history of subfertility, the height of pregnant woman, weight of pregnant woman, BMI of pregnant woman, haemoglobin level, blood sugar, the weight of pregnant woman until six months pregnancy, fundal height and gender of the baby. B. Data collection: When a woman begins her pregnancy *in* Sri Lanka, she is registered for antenatal care, and two sets of records are kept: “A” card and "B" card. During a prenatal clinic appointment, pregnant women hold the "A" card in their hands and display it to health care staff. Midwives keep track of the "B" for record-keeping. Both 500 “A” cards and “B” cards were collected from different MOH divisions areas from Sri Lanka during this study. Record card keeps several records in their sheet. Such as the name of the pregnant woman, age, employments of both wife and husband, BMI, blood group of a pregnant woman, and so on. Selected 16 attributes were extracted from both A and B cards and manually entered. At last baby's birth baby weight grouped into two classes. Baby weight less than 3,200g is considered one class, and baby weight greater than or equal to 3,200g is regarded as another class. Further, from the concerns from the medical practitioners, some factors affecting baby weight *are* grouped into classes according to their problems. Such as, normal BMI should be from 18.5 to 25, average haemoglobin level for pregnant women is 12 to 16, below 12 it's considered as iron deficiency, and below 10 considering as anaemia, regular fasting blood sugar for the pregnant woman is from 90 to 100, and correct age of having pregnancy is 20 to 35. Table II indicates the attributes and the relevant description during data collection of the pregnant woman. Fig. 1. Proposed Approach 429 . . TABLE II. IDENTIFIED ATTRIBUTES AND THE RELEVANT DESCRIPTION No Attribute Description Representation Levels/unique values 1. Gravidity Gravidity is defined as the number of times a woman has been pregnant. Integer E.g., 3 (The pregnant woman has already given birth to 2 children) 2. Age The current age of the pregnant woman Range 0: <20 1: 20-35 2: >35 3. Blood group The blood group of the pregnant woman Varchar A+,A-,B+,B-,O+,O-,AB+,AB- 4. Employment pregnant woman During pregnancy, she worked or not Scale 0: Unemployed 1: Employed 5. History of abortion The pregnant woman experienced with abortion or not Range 0: Not Aborted 1: Aborted 6. Consanguinity Malformations at birth, growth of the mind, sickness, and humanity *are* linked to consanguinity, which can be said as a blood relation's marriage or union. Scale 0:Non Consanguinity 1: Consanguinity 7. Pre-pregnancy scanning The pregnant woman pre-scanned or not Factors 2 level, 0 for not pre-scanned, 1 for pre-scanned 8. History of subfertility Pregnant woman experience in late pregnancy or not through her history since the marriage Scale 0: Non-late pregnancy 1: Late pregnancy 9. Height Height of pregnant woman before pregnancy Float in meters E.g., 1.52 m 10. Weight Weight of pregnant woman before pregnancy Float in kilograms E.g.: 52.5 kg 11. BMI BMI of the pregnant woman before pregnancy Range 23.3, three levels, 0 for below 18.5, 1 for between 18.5 to 25, and 2 for above 25 12. Haemoglobin level Haemoglobin level at the beginning and middle of pregnancy Range 11.1, 3 levels, 0 for below 10, 1 for 10-12, 2 for above 12 13. Blood sugar Blood sugar at the beginning and middle of pregnancy Range 0: <90 1: 90-100 2: >100 14. Weight of pregnant woman from 1 – 6 months The weight measures pregnant women until the first six months Float in kilograms E.g.: 52.5 kg 15. Fundal Height Distance from the pubic bone to the uterus's top Integer in centimeter E.g., 20 cm 16. Gender of the baby Gender of the newborn baby Varchar Girl or Boy C. Data preprocessing Data preprocessing transforms raw data into usable information, then fed into the training model for accurate medical decisions, diagnoses, predictions, and treatments [24]. The data collection set consists of both discrete and continuous variables. In case the uniformity of the values is maintained in the data set. Some attributes have textual format thought the data set. Such as blood group and gender. Therefore, those textual datasets are categorized into groups, and groups are labelled in a numeric way using the scikit-learn library. It's *one* of the ML libraries for python programming. Further, the collection of data set included nan/blank values. In case those nan values changed into numeric values using the NumPy library. D. Data preparation Part of analyzing ML involves splitting data into *training* and testing sets. Once data sets were divided into two groups, the training set was allocated with the majority (80%) of data. On the other hand, other testing sets received a lesser (20%) portion of data. Analysis Services takes a random sample of the data to guarantee *that* *the* *training* *and* testing sets are comparable. We can decrease the effects of data discrepancies and better understand the model's properties by utilizing the same amount of data for training and testing. After preprocessing was carried out, the data set was split into *training* and testing using the ski learn library. In data preparation training data set is trained to apply the ML algorithms in advance. E. Implementation of ML models After completion of data preparation, ML algorithms were used for *the* *training* and testing data set. In this research, five ML algorithms *were* *used* to predict newborn baby weight at the six months of pregnancy. Construction of ANN model: The ANN approach design consists of four layers. Thus, input layer, first hidden layer, second hidden layer, and output layer. The following Table III shows the configuration for the neural network. TABLE III. CONFIGURATION FOR THE NEURAL NETWORK MODEL Parameter Value Epochs 50 Batch size 32 Optimizer Adam Loss Binary_crossentropy Activation (first hidden layer) ReLu Activation (second hidden layer) ReLu Activation (output layer) Sigmoid Construction of SVM: SVM *is* *one* of the supervised learning algorithms *used* *for* classification and regression problems. In this approach, SVM kernel 'RBF' is used for *training* and testing the data set. Further, Logistic Regression, NB, and Decision Tree are applied to *the* *training* and testing data set. IV. RESULTS AND DISCUSSION The research was performed on Windows operating system with Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz 1.99 GHz and 8 GB RAM computer. Python programming 430 language implements preprocessing and the five different ML algorithms. Factors are collected from maternal "A" and "B" records and manually entered into the system. The baby weight data set is classified into two classes to train the ML model. A. Data Preprocessing Data preprocessing is a process to get the useable information to train the data set for the ML model, which was, explained in detail in the previous section. For example, before the preprocessing, the data shows gender as boy and girl. But after its preprocessing, it’s shown as 0 and 1. This is also applicable to the factor of the blood group of pregnant women. B. Performance of ML models To compute the model error, the ANN uses a loss function technique. It's a technique for altering the weights and biases of data input. The optimization procedure updates the weight parameters to decrease the loss function and determine the model error. Fig 2 depicts the model train by epoch. Fig. 2. Model train by epoch The signal processing community developed the Receiver Operating Characteristics (ROC) curve to assess a human operator's ability to differentiate informative radar signals from background noise. It was then primarily utilized in the medical decision-making community to evaluate the utility of a diagnostic test. The trade between the true positive rate and the false-positive rate for a predictive model utilizing different probability thresholds is summarized by ROC Curves. Thus, the following equations explain the true positive and false positive rates in more detail. 4 g, B ñ (1) „ g, B ñ2 ƒ"’ The model's efficiency is judged by measurement precision. The most important criterion for model evaluation is accuracy. The number of tuples correctly classified *based* on the formula is accuracy. The following formula explains the accuracy in detail. 2 â ñ2 ñ2 (3) Fig 3 shows the accuracy of the final ANN model for *the* *training* and testing dataset. Fig. 3. Accuracy of the final ANN model for *the* *training* and testing dataset Further, we use recall, F1 score, precision to evaluate each model using five algorithms. The recall is the method of correctly identifying the true positives. The following formula gives it. Ar QR ñ (4) Simply precision is defined as the ratio between true positives and all the positives, and the following formula defines it. bAr q’ ñ (5) F1 score is a function of precision and recall, and it can be defined in the following formula. 1 rB !"'" $ %&%'() !"" $ %&%'() (6) As from the equations mentioned above, TP indicates true positive, TN indicates true negative, FP indicates false positive, and FN indicates false negative. In this research, TP mentions that if the predicted class is 1, real results are also 1. TN mentions that if the predicted class is 0, real results are also 0. FN mention as, if predicted class 0, but real results is 1. Table IV summarizes each model's recall, precision, and F1 score. Further, Fig 4 shows the comparison of models with percentage accuracy. TABLE IV. RECALL, PRECISION, AND F1 SCORE FOR EACH OF THE IMPLEMENTED MODEL Recall % Precision % F1 score% ANN 73.97 72.97 73.46 SVM 67.74 60.00 63.63 NB 74.19 65.71 69.69 Logistic Regression 65.51 54.28 59.37 Decision Tree 54.54 51.42 52.94 Fig. 4. Accuracy of each model 431 . . . . . . V. CONCLUSION AND FUTURE WORK This research paper identifies the process to predict the newborn baby's weight at six months of pregnancy using ML algorithms. Five different machine algorithms are used for trained data to check *the* effectiveness of prediction mechanisms. In conclusion, ANN and NB models have given accuracy above 70%, SVM and logistic regression algorithms provide accuracy above 60%, and Decision Tree provides accuracy below 60%. The future work of this research can be improved in many ways. We planned to increase the baby weight class to get a more accurate prediction. Also, we are planning to predict the appropriate delivery method for a pregnant woman. Further, we are planning to increase the attribute set too. ACKNOWLEDGEMENT The authors would like to thank the MOH division and medical practitioners in Sri Lanka for their valuable advice and support in collecting *information* on the post-pregnant woman. REFERENCES [1] F. M. Alkhateeb and K. Osias, "MacKinnon III, G. Understanding Health Outcomes and Pharmacoeconomics. Burlington, MA. Jones & Bartlett Learning 2011. 218 pages. $79.95. ISBN 978-0-7637-7099-0," ed: American Journal of Pharmaceutical Education, 2012. [2] N. H. Hillman, S. G. Kallapur, and A. H. Jobe, "Physiology of transition from intrauterine to extrauterine life," Clinics in perinatology, vol. 39, no. 4, pp. 769-783, 2012. [3] S. U. Morton and D. Brodsky, "Fetal physiology and the transition to extrauterine life," Clinics in perinatology, vol. 43, no. 3, pp. 395-407, 2016. [4] S. J. Weiss, P. Wilson, M. J. Hertenstein, and R. Campos, "The tactile context of a mother’s caregiving: implications for attachment of *low* birth weight infants& ," Infant Behavior and Development, vol. 23, no. 1, pp. 91-111, 2000. [5] C. Tiruneh and D. Teshome, "Prediction of Birth Weight by Using Neonatal Anthropometric Parameters at Birth in Finote Selam Hospital, Ethiopia," Pediatric Health, Medicine and Therapeutics, vol. 12, p. 259, 2021. [6] A. G. Philip, G. A. Little, D. R. Polivy, and J. F. Lucey, "Neonatal mortality risk for the eighties: the importance of birth weight/gestational age groups," Pediatrics, vol. 68, no. 1, pp. 122-130, 1981. [7] M. Group. (2021, October 2021). 10 Factors That Affect A Baby's Birth Weight. Available: https://www.marsden-weighing.co.uk/blog/factors- affect-baby-birth-weight [8] C. S. Metgud, V. A. Naik, and M. D. Mallapur, "Factors Affecting Birth Weight of a Newborn–a Community Based Study in Rural Karnataka, India," PloS one, vol. 7, no. 7, p. e40040, 2012. [9] M.-Y. Chang, C.-H. Kuo, and K.-F. Chiang, "The effects of pre- pregnancy body mass index and gestational weight gain on neonatal *birth* weight in Taiwan," International Journal of Nursing and Midwifery, vol. 2, no. 2, pp. 28-34, 2010. [10] W. H. Organization, "Promoting optimal fetal development: report of a technical consultation," 2006. [11] S. M. Marcus, "Depression during pregnancy: rates, risks and consequences," Journal of Population Therapeutics and Clinical Pharmacology, vol. 16, no. 1, 2009. [12] R. Jennifer L.W. Fink, BSN. (2020, October 2021). 8 Health Factors That Affect a Baby's Birth Weight. Available: https://www.healthgrades.com/right-care/pregnancy/8-health-factors- that-affect-a-babys-birth-weight [13] T. L. Gross, R. J. Sokol, T. Williams, and K. Thompson, "Shoulder dystocia: a fetal-physician risk," American journal of obstetrics and gynecology, vol. 156, no. 6, pp. 1408-1418, 1987. [14] J. B. Whitfield, S. A.

Treloar, G. Zhu, and N. G. Martin, "Genetic and non-genetic factors affecting birth-weight and adult body mass index," Twin Research and Human Genetics, vol. 4, no. 5, pp. 365-370, 2001. [15] E. Dougeni et al., "Dose and image quality optimization in neonatal radiography," The British journal of radiology, vol. 80, no. 958, pp. 807-815, 2007. [16] N. W. Svenningsen, "Neonatal Intensive Care: When and Where Is It Justified?," International Journal of Technology Assessment in Health Care, vol. 8, no. 3, pp. 457-468, 1992. [17] Y. Lu, X. Zhang, X. Fu, F. Chen, and K. K. Wong, "Ensemble *machine* learning for estimating *fetal* weight at varying gestational age," in Proceedings of the AAAI conference on artificial intelligence, 2019, vol. 33, no. 01, pp. 9522-9527. [18] D. Dissanayake, D. Vidanagama, and N. Wedasinghe, "Web-Based Workload Maintain System for Midwives in Sri Lanka," 2020. [19] J. D. Boardman, D. A. Powers, Y. C. Padilla, and R. A. Hummer, "Low birth weight, social factors, and developmental outcomes among children in the United States," Demography, vol. 39, no. 2, pp. 353-368, 2002. [20] J. V. De Bernabé et al., "Risk factors for low birth weight: a review," European Journal of Obstetrics & Gynecology and Reproductive Biology, vol. 116, no. 1, pp. 3-15, 2004. [21] A. L. Beam and I. S. Kohane, "Big data and *machine* learning in health care," Jama, vol. 319, no. 13, pp. 1317-1318, 2018. [22] B. N. E. Morton, "The inheritance of human birth weight," Annals of Human Genetics, vol. 20, no. 2, pp. 125-134, 1955. [23] A. Talie, M. Taddele, and M. Alemayehu, "Magnitude of *low* *birth* weight and associated factors among newborns delivered in Dangla primary hospital, Amhara regional state, Northwest Ethiopia, 2017," Journal of Pregnancy, vol. 2019, 2019. [24] K. Al-Jabery, T. Obafemi-Ajayi, G. Olbricht, and D. Wunsch, Computational Learning Approaches to Data Analytics in Biomedical Applications. Academic Press, 2019. [25] S. Karthiga, K. Indira, and C. N. Angeline, "Machine Learning Model to Predict Birth Weight of New Born Using TensorFlow," Computer Science & Information Technology, pp. 71-90, 2019. [26] J. Tao, Z. Yuan, L. Sun, K. Yu, and Z. Zhang, "Fetal birthweight prediction with measured data by a temporal machine learning method," BMC Medical Informatics and Decision Making, vol. 21, no. 1, pp. 1-10, 2021. [27] Z. Hussain and M. D. Borah, "Birth weight prediction of new born baby with application *of* *machine* learning techniques on features of mother," Journal of Statistics and Management Systems, vol. 23, no. 6, pp. 1079-1091, 2020. [28] M. O. Al-Shawwa and S. S. Abu-Naser, "Predicting birth weight using artificial neural network," 2019. .