International Journal of Scientific and Innovative Mathematical Research(IJSIMR)
Volume 11, Issue 2, 2023, PP 32-39
ISSN No. (Online) 2347-3142
https://doi.org/10.20431/2347-3142.1102003
www.arcjournals.org

Prediction of Pre-Pregnancy Women and Infant Birth Weight Gain through Machine Learning among Mothers Receiving Gynecological Care

K. Sujatha, D, Flora Evangil, G. Manimannan

Assistant Professor, Department of Mathematics, ST. Joseph's College (Arts & Science), Kovur, Chennai

Corresponding Author:** ***K. Sujatha, D, *Assistant Professor, Department of Mathematics, ST. Joseph's College (Arts & Science), Kovur, Chennai*

Abstract: *This research paper looks into the use of Artificial Intelligence (AI) and Machine Learning (ML) methods in the field of obstetrics and gynecology. Specifically, a ML model is proposed to predict weight gain events during pregnancy based on clinical parameters such as education, age, height, pre-pregnancy weight, pre-pregnancy BMI, ninth month weight, weight gain, baby weight, neonatal complication, Apgar score and LSCS/ND.This study collected data from 50 pregnant women in Chennai, including clinical parameters such as education, age, weight, and baby weight, etc.. The database was split into three classes and used to evaluate four machine learning algorithms. The logistic regression model achieved the highest performance metrics and is the most effective for predicting birth weight. However, the other models performed well too. Using a heat map to visualize the data helped to detect correlations between different variables. The study concludes that machine learning models can be useful tools for predicting birth weight and visualizing the data with a heat map can provide valuable insights for further analysis.*

Keywords: *Machine Learning Models, Logistic Regression, Confusion Matrix, Heat Map and Visualization*

1. Introduction

Pregnancy weight gain is a key indicator of maternal nutrition and intra-uterine fetal nutrition, and sub-optimal Gestational Weight Gain (GWG) can lead to a range of adverse outcomes, such as High Birth Weight (HBW), Low Birth Weight (LBW), pregnancy-induced hypertension, gestational diabetes, preterm births, caesarean delivery, and delayed initiation of breastfeeding [1–5]. Birth weight (BW) is a predictor of fetal wellbeing, newborn survival, and is heavily dependent on maternal health and nutrition during pregnancy [6]. Growth failure in children is most likely to occur in the critical window of opportunity, from conception up-to two years of age, and around half of growth failure which occurs by two years of age occurs in uterus [7].

The long-term consequences of intrauterine malnutrition are serious and far reaching, and can permanently affect the structure and function of tissues [8]. To break the cycle of malnutrition, the fetal period is a critical window of opportunity for nutrition intervention and improving birth weight [7]. Although there are various proven interventions for pregnant women, such as weight monitoring, health education, and counseling on weight management, nutrition and physical activity, unfortunately the majority of pregnant mothers in developing countries do not have access to this information and these services [9]. This has caused pregnancy weight gain to remain a low priority for antenatal care providers and pregnant mothers in developing countries [10]. In 2009, the Institute of Medicine published a weight gain guideline for singleton pregnancy based on pre-pregnancy body mass index classes.

This study aimed to determine the effect of pregnancy weight gain on birth weight among pregnant mothers receiving antenatal care from private clinics in Chennai, India. This is especially pertinent given that most previous studies have been conducted in developed countries and there is a lack of information regarding the effect of pregnancy weight gain on birth weight in developing countries [1, 9]. Pregnancy weight gain, regardless of pre-pregnancy body mass index (BMI), is an independent predictor and has been found to have a significant effect on fetal growth [1, 6]. For underweight mothers, a recommended gestational weight gain (GWG) is 12.5–18 kg; for normal weight, 11.5–16 kg; for overweight, 7–11.5 kg; and for obese mothers, 5–9 kg [9].

2. Material and Methods

2.1. Database

Data was collected from 50 pregnant women in a private hospital in Chennai. The database included clinical parameters such as education, age, height, pre-pregnancy weight, pre-pregnancy BMI, ninth month weight, weight gain, baby weight, neonatal complication, Abgar score, and LSCS/ND. It was further divided into three classes: High, Low, and Moderate Baby Weight Gain.

2.2. Logistic Regression (LR)

Logistic regression in Orange Data Mining is a supervised machine learning algorithm used for classification tasks. It is used to predict the probability of an event occurring based on the values of independent variables. It is a type of regression analysis that is used to predict a categorical dependent variable, such as whether a customer will purchase a product or not. Logistic regression is a powerful tool for predicting the probability of an event occurring, and it can be used to identify important factors that influence the outcome.

Algorithm of Logistic Regression in Orange Data Mining

Step 1. Load the data set into Orange Data Mining.

Step 2. Select the Logistic Regression algorithm from the list of available algorithms.

Step 3. Select the target variable and the predictor variables.

Step 4. Set the parameters for the algorithm, such as the regularization parameter, the maximum number of iterations, and the learning rate.

Step 5. Run the algorithm and view the results.

Step 6. Evaluate the model performance using the confusion matrix, ROC curve, and other metrics.

2.3. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm used for classification and regression tasks. It is a powerful and versatile algorithm that can be used to solve a wide variety of problems. In Orange Data Mining, SVM is used to build predictive models from data. It works by mapping data to a high-dimensional feature space so that data points can be categorized, even when the data are not otherwise linearly separable. SVM is a powerful tool for data analysis and can be used to build predictive models from data. The following algorithm is to use for SVM learner model.

Step 1. Load the dataset and split it into training and testing sets.

Step 2. Pre-process the data by normalizing or standardizing the features.

Step 3. Choose a kernel function and set the hyper parameters.

Step 4. Train the SVM model using the training set.

Step 5. Test the model using the testing set and measure the accuracy.

Step 6. Tune the hyper parameters to improve the accuracy.

Step 7. Evaluate the model using cross-validation and other metrics.

Step 8. Make predictions on new data using the trained model.

2.4. k-Nearest Neighbor (k-NN)

k-Nearest Neighbor (k-NN) is a supervised machine learning algorithm used for classification and regression. It is a non-parametric method used for classification and regression. In k-NN, the data points are classified based on their similarity to other data points. The data points are assigned to a class based on the majority of the k-nearest data points. k-NN is a simple and effective technique for classification and regression problems. It is used in many applications such as recommendation systems, image recognition, and anomaly detection.k-NN algorithm is discussed in the following steps.

Step 1. Calculate the distance between the query point and all the points in the dataset.

Step 2. Select the k-nearest points from the dataset.

Step 3. Calculate the average of the k-nearest points.

Step 4. Assign the query point to the class with the highest average.

2.5. AdaBoost

AdaBoost (Adaptive Boosting) is a supervised machine learning algorithm used for classification and regression problems in Orange Data Mining. It is a meta-algorithm that combines multiple weak learners to form a strong learner. It works by sequentially training each weak learner with a weighted version of the data, where more weight is given to examples that were misclassified by the previous weak learners. The final model is a weighted combination of all the weak learners. AdaBoost is a popular algorithm because it is simple to implement, computationally efficient, and often produces good results. AdaBoost algorithm is explained in the following steps.

Step 1. Initialize the weights of all training examples to be equal.

Step 2.For each iteration of the algorithm:

Step 2.1. Train a weak classifier using the weighted training examples.

Step 2.2. Calculate the error rate of the weak classifier.

Step 2.3. Calculate the weight of the weak classifier.

Step 2.4. Update the weights of the training examples.

Step 3. Combine the weak classifiers to form a strong classifier.

2.6. Test and Score

Test and Score in Orange Data Mining is used to evaluate the performance of a predictive model. It allows users to compare different models and select the best one for their data. It also provides a way to measure the accuracy of a model and identify areas for improvement. Test and Score can be used to compare different algorithms, such as decision trees, neural networks, and support vector machines. It can also be used to compare different parameter settings for a given algorithm.

2.7. Confusion Matrix

A confusion matrix is a table that is used to evaluate the performance of a classification model. It is a table of the predicted class labels and the actual class labels. It allows you to see where the model is making correct predictions and where it is making incorrect predictions. It is a useful tool for evaluating the accuracy of a model and for identifying potential areas of improvement. In Orange Data Mining, confusion matrices can be used to evaluate the performance of classification models such as decision trees, logistic regression, and support vector machines.

2.8. Heat maps

Heat maps are a powerful tool in Orange Data Mining for visualizing data. They allow users to quickly identify patterns and correlations in their data, as well as to identify outliers and clusters. Heat maps can be used to explore relationships between variables, to compare different datasets, and to identify areas of interest. Heat maps can also be used to identify areas of potential improvement in data analysis and machine learning models.

2.9. Model Training

Learning from multi-label data can be accomplished through various methods, such as data transformation, adapting traditional classification models, and the use of ensembles of classifiers [11]. The data transformation approach involves transforming the original multi-label data into one or more binary or multiclass datasets. Adaptation of existing models, on the other hand, entails modifying existing classification algorithms such that they can process multi-label data and produce multiple outputs instead of one [11]. Some models that are suitable for multi-label classification are Decision Tree Classifier, Extra Tree Classifier, Random Forest Classifier, and K-Nearest Neighbors Classifier. Decision Tree Classifier is a non-parametric supervised learning algorithm that can be used to solve classification tasks. It is represented by a tree structure consisting of a root node, branches, internal nodes, and leaf nodes. It works using the divide-and-conquer strategy, which is a recursive partitioning of the problem into smaller sub-problems until it becomes easy enough to be solved directly.

The k-Nearest Neighbors (k-NN) Classifier is an instance-based learning algorithm that is lazy in its induction process. KNN assumes that instances with similar properties are located close to each other. It works by calculating the distance between an unclassified instance and all the other instances in the dataset and then selecting the - closest ones. To calculate the distance, the Euclidean distance was used and the k-value chosen was five. Two methods from the literature were applied for data transformation: Binary Relevance and Label Power set. Binary Relevance is a method of decomposing the learning of each label into a set of binary classification tasks, one per label, in which each model is independently learned from the information from that label and ignoring the information from the other labels [12].

The Label Power Set (LPS) method is an approach for dealing with multiclass classification problems, wherein each different combination of labels is used as an identifier of a new class [12]. As a result, the dataset is reduced to a single class, making it easier to treat as a multiclass classification [14]. This technique has the advantage that any binary learning algorithm can be used as an estimator, and it has linear complexity for the number of labels. In this research paper, Logistic Regression, Support Vector Machine, K-Neighbors Classifier, and AdaBoost Classifier were used as estimators. A training set was used to train all these models, and the performance of the models was evaluated by testing the LPS method. One drawback of this technique is that it does not take into consideration any label dependency.

2.10. Model Validation

After training the models, we use the test set to evaluate their performance. Model selection was assessed by calculating AUC ROC (area under the curve ROC), accuracy (fraction of instances that the model classified correctly), precision (proportion of positive identifications that were actually positive), recall (proportion of the positive class that was correctly classified), F1 score (harmonic mean of precision and recall), and Hamming loss (proportion of misclassifications). Two methods of label-based measures were used: macro-average and micro-average. Macro-average computes the metric independently for each label, giving equal weight to all labels. On the other hand, micro-average metrics aggregate the contributions of all labels to compute the average metric [13].

$$Accuracy = \frac{TP}{TP+TN+FP+FN} \tag{1}$$

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$F1\ Score = \frac{2*TP}{2*TP+FP+FN} \tag{4}$$

3. Result and Discussion

In this study, the database was split into 20 percent for training and 80 percent for testing. To use the entire ML algorithm, the best model was achieved and its AUC, F1, CA, Precision and Recall scores were summarized in Table 1. This table presents the performance metrics of four different machine learning models - Logistic Regression, SVM Learner, k-NN, and AdaBoost, evaluated on a binary classification task. The evaluation metrics used to assess the performance of the models are AUC, CA, F1 score, precision, and recall.

AUC (Area Under the Curve) is a metric that represents the performance of the model in terms of its ability to distinguish between the positive and negative classes. It measures the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate against the false positive rate. A higher AUC score indicates better performance of the model.

CA (Classification Accuracy) is the proportion of correctly classified instances out of the total instances in the dataset. It measures the overall accuracy of the model in predicting the class labels.

F1 score is a harmonic mean of precision and recall. It is a measure of the model's ability to correctly identify both the positive and negative classes.

Precision is the proportion of true positives out of the total predicted positives. It measures the accuracy of the positive predictions made by the model.

Recall is the proportion of true positives out of the total actual positives. It measures the model's ability to correctly identify the positive instances.

From the table, it can be observed that the Logistic Regression model performs the best among all the models with an AUC score of 0.99. The model has high precision and recall, indicating that it can correctly identify both the positive and negative classes. The model also has a high classification accuracy of 0.95, indicating that it can accurately classify the instances.

The other models, namely SVM Learner, k-NN, and AdaBoost, also perform well with AUC scores of 0.98 and F1 scores of 0.93-0.95. However, the classification accuracy of these models is slightly lower than that of the Logistic Regression model.

Overall, the table suggests that the Logistic Regression model is the best performing model for the given classification task, followed by SVM Learner, k-NN, and AdaBoost (Table 1).

Table1. *ML Algorithm Summary Statistics*

Model	AUC	CA	F1	Precision	Recall
Logistic Regression	0.99	0.95	0.95	0.95	0.95
SVM Learner	0.98	0.93	0.93	0.93	0.93
k-NN	0.98	0.94	0.94	0.94	0.94
AdaBoost	0.96	0.95	0.95	0.95	0.95

This (Table 2) shows the confusion matrix for a multi-class classification problem where the objective is to predict the birth weight of babies based on the gestational weight gain of their mothers. The actual class labels are represented in the rows of the table, and the predicted class labels are represented in the columns of the table. The table shows the performance of four different machine learning models, namely Logistic Regression, SVM, k-NN, and AdaBoost.

The confusion matrix table is divided into four rows and four columns, with a total of 16 cells. Each cell represents the number of instances that fall into a particular combination of the actual and predicted classes.

The first row represents the actual Gestational Weight Gain (GWG) class, where there were a total of 9 instances. Out of these 9 instances, the Logistic Regression model predicted 8 instances as GWG and 2 instances as Low Birth Weight. SVM model predicted 8 instances as GWG and 2 instances as Low Birth Weight. k-NN model predicted 8 instances as GWG and 1 instance as Low Birth Weight. AdaBoost model predicted 8 instances as GWG and 1 instance as Low Birth Weight. This means that all the models have correctly predicted 8 instances as GWG, while the Logistic Regression, SVM, and AdaBoost models misclassified 2 instances as Low Birth Weight, and k-NN misclassified 1 instance as Low Birth Weight.

The second row represents the actual High Birth Weight (HBW) class, where there were a total of 25 instances. Out of these 25 instances, all the four models predicted correctly all 25 instances as HBW. This means that all the models have correctly classified all the instances as HBW.

The third row represents the actual Low Birth Weight (LBW) class, where there were a total of 16 instances. Out of these 16 instances, the Logistic Regression model predicted 1 instance as GWG and 14 instances as LBW. SVM model predicted 2 instances as GWG and 14 instances as LBW. k-NN model predicted 2 instances as GWG and 14 instances as LBW. AdaBoost model predicted 1 instance as GWG and 14 instances as LBW. This means that all the models have correctly predicted 14 instances as LBW, while the Logistic Regression and AdaBoost models misclassified 1 instance as GWG, and SVM and k-NN models misclassified 2 instances as GWG.

The fourth row represents the sum of the instances for each column. Each column shows the total number of instances predicted as GWG, HBW, and LBW. The sum of the instances for each column matches the total number of instances in the dataset, which is 50.

In summary, the confusion matrix table provides a comprehensive evaluation of the performance of the machine learning models. It shows the number of true positives, true negatives, false positives, and false negatives for each class, allowing us to analyze the strengths and weaknesses of each model. Based on the confusion matrix table, it can be concluded that all the models performed well in predicting the HBW class, while the Logistic Regression and AdaBoost models performed better than SVM and k-NN models in predicting the GWG and LBW classes.

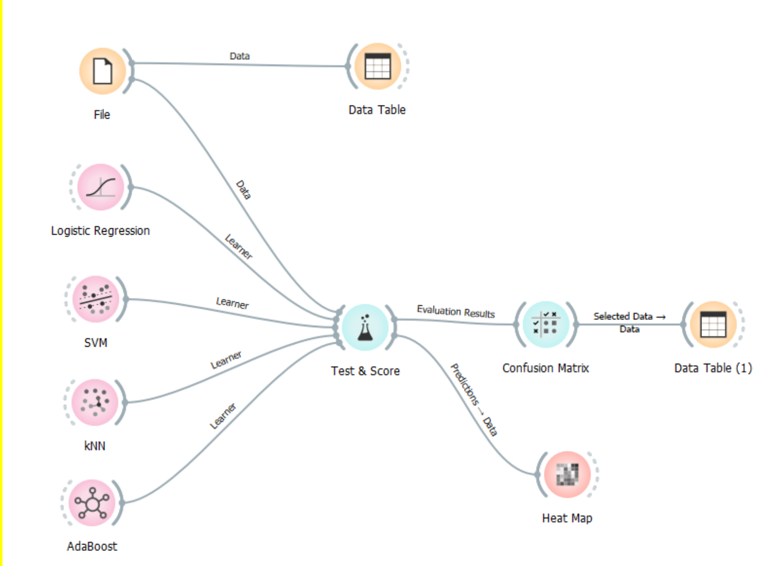

Figure1. *Work Flow Diagram for ML Algorithms*

Table2. *Confusion Matrix of ML Algorithms (Baby Birth Wight gain)*

MLA		**Predicted**				
			Gestational Weight Gain	High Birth Weight	Low Birth Weight	Sum
Logistic Regression	**Actual**	Gestational Weight Gain	8	0	2	10
		High Birth Weight	0	25	0	25
		Low Birth Weight	1	0	14	15
		Sum	9	25	16	50
SVM	**Actual**	Gestational Weight Gain	8	0	2	10
		High Birth Weight	0	25	0	25
		Low Birth Weight	1	0	14	15
		Sum	9	25	16	50
k-NN	**Actual**	Gestational Weight Gain	8	0	1	9
		High Birth Weight	0	24	1	25
		Low Birth Weight	0	2	14	16
		Sum	8	26	16	50
AdaVoost	**Actual**	Gestational Weight Gain	8	0	1	9
		High Birth Weight	0	25	0	25
		Low Birth Weight	2	0	14	16
		Sum	10	25	15	50

Figure 2 in the Orange data mining tool presents a heat map, which offers a visual representation of the data that enables users to identify patterns and correlations between variables easily. Heat maps can detect similar data points, outliers, and relationships between variables. Additionally, they can be used to locate areas of interest, such as regions with high or low correlation, or areas with high or low density. This research paper visualized three clusters of four machine learning algorithms models using the heat map. The first cluster comprises weight gain, baby weight, education, and LSCS/ND.

The second cluster consists of pre-pregnancy BMI, pre-pregnancy weight, and ninth-month pregnancy weight. Finally, the third cluster includes age, Agbar score, and neonatal complication.

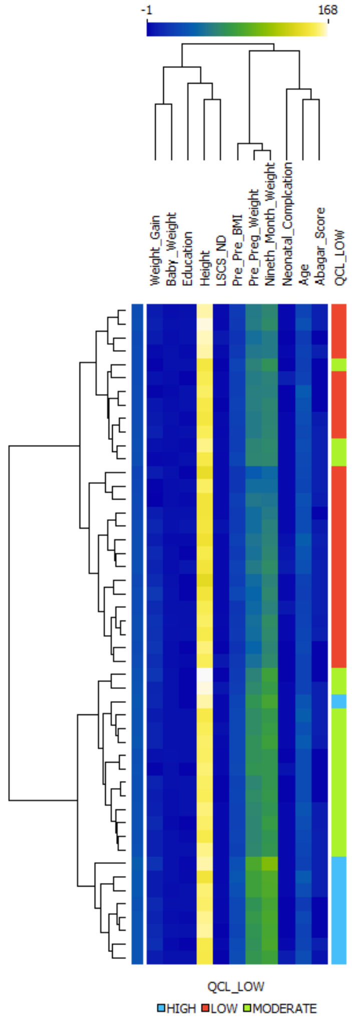

Figure2. *Heat Map of GWG, HBW and LBW (High=GWF, LOW= HBW and Moderate=LBW)*

4. Conclusion

In this study, split the database into training and testing sets, and evaluated the performance of four machine learning algorithms - Logistic Regression, SVM Learner, k-NN, and AdaBoost - on binary and multi-class classification tasks.. Firstly, all of the models achieved high accuracy, with AUC scores ranging from 0.96 to 0.99. This indicates that the models are reliable and effective in predicting birth weight. Secondly, the logistic regression model achieved the highest performance metrics, with a F1 score, precision, and recall of 0.95. This indicates that logistic regression is the most effective model for predicting birth weight in our dataset. However, the other models, including SVM, k-NN, and AdaBoost, also performed well and can be considered as viable options for predicting birth weight. Finally, using a heat map to visualize the data helped to identify clusters of similar data points and to detect correlations between different variables. Overall, our findings suggest that machine learning models can be useful tools for predicting birth weight, and that visualizing the data with a heat map can provide valuable insights for further analysis.Top of Form

References

[1] Hector D, Hebden L. Prevention of excessive gestational weight gain: An evidence review to inform policy and practice. Phys Act NutrObes Res Gr. 2013; (February).

[2] Muthayya S. Maternal nutrition & low birth weight—What is really important? Indian J Med Res. 2009; 130(5):600–8. PMID: 20090114

[3] Panahandeh Z. Gestational Weight Gain and Fetal Birth Weight in Rural Regions of Rasht / Iran. 2009; 19(1):18–24.

[4] Liu Y, Dai W, Dai X. Pre-pregnancy body mass index and gestational weight gain with the outcome of pregnancy: a 13-year study of 292, 568 cases in China. 2012; 905–11.

[5] Haugen M, Brantsæter A, Winkvist A, Lissner L, Alexander J, Oftedal B, *et al.* Associations of pre-pregnancy body mass index and gestational weight gain with pregnancy outcome and postpartum weight retention: a prospective observational cohort study. BMC Pregnancy Childbirth. 2014; 14(1):201.

[6] Lechtig A, Cornale G, Ugaz ME, Arias L. Decreasing stunting, anemia, and vitamin A deficiency in Peru: Results of the good start in life program. Food Nutr Bull. 2009; 30(1):37–48. https://doi.org/10.1177/156482650903000104 PMID: 19445258

[7] Shrimpton R. Maternal Nutrition and the Intergenerational cycle of growth failure. Matern Nutintergenerational cycle growth Fail ProgNutr 6th Rep World NutrSituat. 2010; 62–75.

[8] Barker DJP. Inutero programming of cardiovascular disease.Theriogenology. 2000; 53(2):555–74. PMID: 10735050

[9] Rasmussen KM, Yaktine AL. Weight Gain during Pregnancy: Reexamining the Guidelines, 2009.

[10] Rasmussen KM, Catalano PM, Yaktine AL. New guidelines for weight gain during pregnancy: what obstetrician / gynecologists should know. 2009.

[11] Kuhn, M.; Johnson, K. Applied Predictive Modeling; Springer Science & Business Media: New York, NY, USA, 2013.

[12] Herrera, F.; Charte, F.; Rivera, A.J.; del Jesus, M.J. Multilabel Classification. In Multilabel Classification: Problem Analysis, Metrics and Techniques; Springer: Cham, Switzerland, 2016; pp. 17–31.

[13] Luaces, O.; Díez, J.; Barranquero, J.; del Coz, J.J.; Bahamonde, A. Binary relevance efficacy for Multilabel classification. Prog.Artif.Intell. 2012, 1, 303–313.

[14] Boutell, M.R.; Luo, J.; Shen, X.; Brown, C.M. Learning multi-label scene classification. Pattern Recognit. 2004, 37, 1757–1771.

Citation: *K. Sujatha, D et al. "Prediction of Pre-Pregnancy Women and Infant Birth Weight Gain through Machine Learning among Mothers Receiving Gynecological Care" International Journal of Scientific and Innovative Mathematical Research (IJSIMR), vol 11, no. 2, 2023, pp. 32-39. DOI:* https://doi.org/*10.20431/2347- 3142.1102003.*