Indian Journal of Public Health Research and Development
Volume 14 No. 2, April-June 2023

Impact of Education and Socio-Economic Status on Post-Natal Body Weight using Machine Learning Approach

Saishree Sahu[1], Prasanna Kumar Dixit[2], Chinmayee Dora[3]

[1]Research Scholar, Department of Zoology, [2]Associate Professor, Department of Zoology, Berhampur University, Odisha, [3]Assistant Professor, Department of ECE, CUTM, Odisha, India

How to cite this article: Saishree Sahu, Prasanna Kumar Dixit, Chinmayee Dora. Impact of Education and Socio-Economic Status on Post-Natal Body Weight using Machine Learning Approach. Indian Journal of Public Health Research and Development 2023;14(2).

Abstract

Background: The maternal education and socio-economic status of the mother greatly affect the growth of the baby. Nutritional deficiency is the major cause of motor and physiological disorders in children. Hence, predicting body weight could be used to monitor the growth of the babies. Although the prediction of weights is standardized by World Health Organization (WHO), for babies based on their gender and age; generalizing it for all locations and socio-economic variations is not possible

Method: Hence, in this paper, an ANN-based predictive approach to establish the relationship between the factors like socioeconomic status, age, and prior weights to predict the baby's weight using a locally acquired dataset. This could be helpful for the health workers to properly diagnose the disease a time ahead.

Conclusion: The proposed work suggests predicting the baby's growth rate using Machine Learning (ML) techniques is both an efficient and feasible approach.

Keywords: Neonatal weight, Socioeconomic status, Machine learning, ANN, Regression.

Introduction

During the prenatal phase, fetal growth is influenced by the mother's nutritional status. Economic, social, and cultural considerations could make it difficult for many women in developing countries to receive the quality food and health care. After the childbirth, baby's physical and intellectual growth solely depends on its nutritional intake and indicated by their body weight. Lack of which malnutrition can occur. Malnutrition could be poor diet or insufficient calorie intake, an uneven calorie intake, and problems induced by sickness, leading to impaired nutrient absorption and excessive nutrient loss, are only a few of them[1]. Quality food is the primary source of nourishment as it provides vital chemicals for development lack of which risk of malnutrition increases[2]. 80% of brain development and organ growth occurs within the two primary years of life, hence early detection of achild's sensitivity to malnutrition is critical[3]. Excessive muscle loss, greater infection, higher complication, increased morbidity and mortality, and poor wound

Corresponding Author: Chinmayee Dora, Assistant Professor, Department of ECE, CUTM, Odisha, India.

Email: chinmayee@cutm.ac.in

healing are all consequences of malnutrition[4]. India has a high population of malnourished children and adults, according to a World Health Organization (WHO) research[5]. The Sustainable Development Goals (SDGs) of eradicating all types ofmalnutrition by 2030, according to a 2018 WHO and UNICEF assessment, are ambitious, but maynot be realized[6]. A curative intervention might be early diagnosis of vulnerability and provisionof needed therapy. Several studies imply that the mother's socioeconomic status has an impact on the child's development. The growth of the child could be quantified using his/her observed body weights over a timespan. So, those factors could be considered for applying Machine Learning (ML) techniques to predict the child's weight in advance. The mean body weight of boy and girl according to WHO standards are presented in the figure 1.

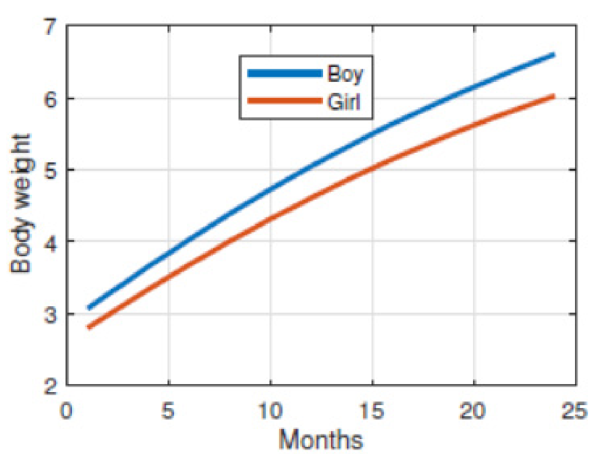


Fig. 1 Mean body weight of boy and girl according toWHO standards

ML techniques might provide a cost-effective solution in terms of money and time for early detection of retarded growth of a child. ML approaches may predict a child's nutritional status with a high degree of accuracy if the model is trained with reliable data.

In this work, ML techniques wereused to achieve high-accuracy predictions based on a few critical characteristics. Because this research focuses on predicting baby weight, the mother's socioeconomic position might be a key starting point.

ML based Body Weight Prediction

From several studies suggested that the child's body weight is in linear relation to the mother's socioeconomic status and qualitative nutritional intakes. So, the present study is to predict child's body weight in-relation-to past record of body weight and socioeconomic status of family. In order to identify a correlation between them, ML algorithm is adopted and implemented. Although ML algorithms are extensively applied for regression and classification purposes such as fiscal analysis[7], remote sensing[8,9], health related like diabetes prediction[10,11]. While, their applications for body weight prediction of child are sparse in literature. Linear and non-linear regression analysis for prediction of neonatal weight at birth is presented in[12]. Several such research has been carried out to predict infant weight using probabilistic models[13], nutritional prediction using artificial intelligence[14], birth weight prediction using neural network[15,16], dog body weight prediction using morphological parameters[17]. Other than data mining, image processing-based body weight prediction also proved efficient[18]. Detection and classification of malnutrition has been successfully implemented using decision tree and artificial neural network approaches[19]. In this study, the child's body weight is predicted using Decision Tree and Artificial Neural Network, and compared to find the best alternative.

Decision Tree

For classification and regression, Decision Trees (DT) could be used as a non-parametric supervised learning method. There are three types of nodes in a decision tree which are decision nodes, chance nodes, and end nodes. The contents of the leaf node are the outcome, and the conditions along the route create a conjunction in the if clause in the decision tree. If-then rules are used to apply to DT in general. Classification or regression rules are represented by the pathways from root to leaf.

Artificial Neural Network

Artificial Neural Networks (ANN) are a type of human brain replication that uses mathematical models to produce a computer method for information processing[20]. ANN is made up of numerous layers with linked neurons, with each neuron represented as shown in Fig.2. Independent variables characterise the input layer, whereas dependent variables characterise the output layer. The ANN could build the association between input and output hyper-dimensional spaces based on the training data. The back propagation algorithm is commonly used in the ANN training process, which reduce the error function.

The ANN is a function of the input (observation) space X and the output (decision) space Y (Eq. 1).

$$f : X \rightarrow Y \tag{1}$$

The linear connection between the neuron's output and its input feature space with features as might be characterised as in Eq. 2.

$$Y' = \sum_{i=1}^{n} x_i w_i + b \tag{2}$$

where is a bias term that is multiplied by the input feature space and added to the sum of weights.

In most real-time applications, highly non- linear systems may be modelled using ANN by including the transfer function. Such a transfer function is characterised by the activation function, as shown in Eq. 3.

$$Y = \varphi(Y') = \varphi\left(\sum_{i=1}^{n} x_i w_1 + b\right) \tag{3}$$

where the activation function is used φ (.) often refers to nonlinear functions[21] like logsig(), tansig(), purelin(), and so on.

Proposed ANN Architecture

The proposed ANN architecture consists of five input features, multiple hidden layers, and a single output neuron for predicting the post-natal body weight. Past three months body weight along with gender, and socioecenomic status of the mother are inclusive of the five input features. The socioeconomic status consists of information such as education, occupation and yearly income of the family. A framework of proposed ANN model is presented in the Figure 2. In the figure, represents the post-natal body weight at present month, similarly and   represents the measured body weights in the past months. represents the future prediction of body weight. The scores representing education, occupation and income as specified in the Table 1 are calculated and fed to the ANN with reference to the collected survey data. So, all the input features are of numeric data except gender which is a binary feature. The ANN model further is trained using back propagation technique based on the training data.

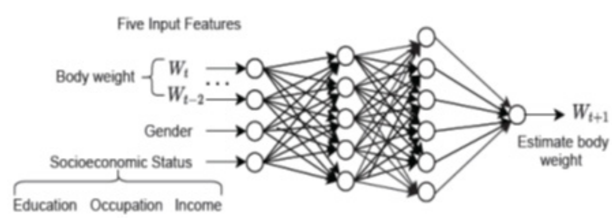


Fig. 2 Proposed framework of Artificial Neural Network

Table 1 Modified-Kuppuswamy scale of reference (February 2019)

Education of Head of the Family	Score	Occupation of Head of the Family	Score	Total Family Income per Month	Score
Professional Degree	7	Professional	10	52,734 and above	12
Graduate or postgraduate	6	Semi professional	6	26,355-52,733	10
Intermediate/diploma	5	Clerical/shop/farm	5	19,759-26,354	6
High school	4	Skilled worker	4	13,161-19,758	4
Middle school	3	Semi-skilled worker	3	7,887-13,160	3
Primary school	2	Un-skilled worker	2	2,641-7,886	2
Illiterate	1	Unemployed	1	Less than 2,640	1

Materials and Methods

Study Area

The study was carried out in various areas of the Ganjam region of Odisha, with the mothers of the volunteers being chosen at random. Ganjam district is one of Odisha's most populous and advanced districts, with inhabitants from various socioeconomic groups. Families from various socioeconomic groups were selected and urban sectors. Data on study parameters was gathered from the sampled families, health workers, community health centres, Anganwadi workers, and the Chief District Medical Officer's office (CDMO). Volunteer families communicated by visitation or over the phone at fixed intervals to acquire data on various aspects of the baby's postnatal development.

Collection of Data

Data was collected throughout and after pregnancy by planned house visits, Anganwadi visits, and phone calls. 312 pregnant women were enrolled in this study, and they were followed up on.

Out of 312 individuals, 300 mothers and their infants (158 boys and 142 girls) were chosen for this study. The research was unable to enrol the 12 individuals since they were unavailable throughout the long-term data gathering phase. According to their socioeconomic position, the women were divided into five categories.

Procedures

Participants were given an information consent document outlining the study's needs and their rights before being asked to submit their information. Pregnant women are asked to provide information such as their age, pregnancy stage, projected delivery date, education, employment, and family monthly income during the initial examination. Following that, post-birth data such as 1) gender, 2) date of birth, 3) birth weight, and 4) health condition were gathered.

Assessment of Mother's Socioeconomic Status (SES)

The SES of women is calculated using the collected data, which includes 1) education, 2) occupation, and 3) family income. According to the Kuppuswamy scale[22] of reference, the SES is further split into five groups. As demonstrated in Table 1 and Table 2, this scale categorizes the sample populations into five SES categories.

Table 2 Socio-Economic-Status (SES), class definitions as per scoring using Modified-Kuppuswamy scale

SES	Class	Total Score
1	Upper	26--29
2	Upper-Middle	16- 25
3	Lower-Middle_3	11--15
4	Upper-Lower	5--10
5	Lower	3--9

Performance Measure

The prediction of post-natal body weight using decision tree approach is validated using three performance measures such as Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE).

The equation for could be represented

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i^{'})^2 \tag{4}$$

The equation for could be represented as,

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_i')^2} \tag{5}$$

The equation for MAE could be represented as,

$$MAE = \frac{1}{n}\sum_{i=1}^{n} Y_i - Y_i' \tag{6}$$

Results and Discussion

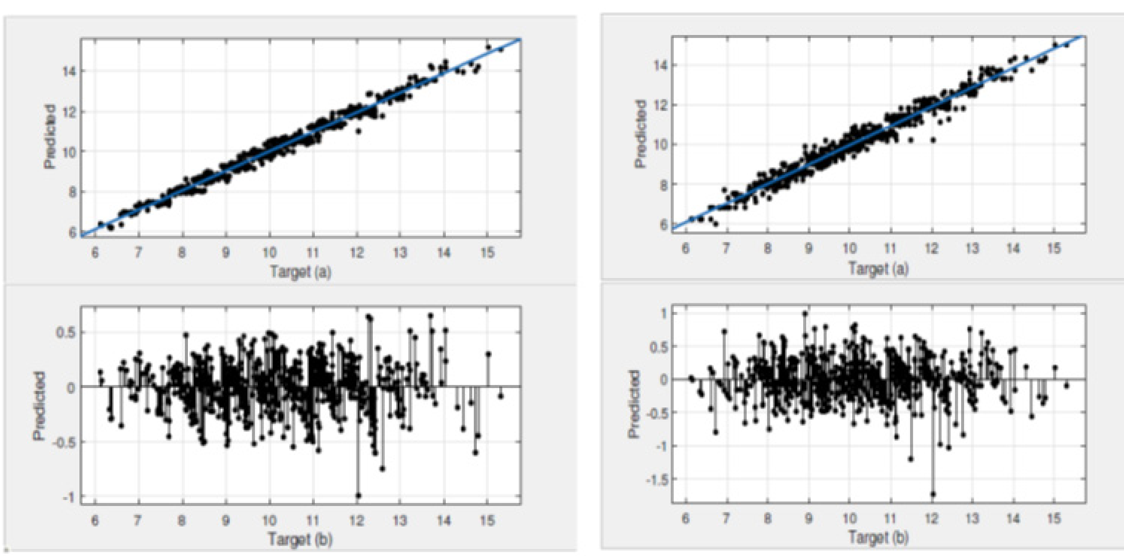
Experiments on the collected data is carried out using both Decision Tree Regressor and Artificial Neural Network Regressor. The collected dataset is divided into 10% training and 90% testing data. All the performances reported for the experiments on the testing data. Performance measures such as and are evaluated for each category of experiments. The results are represented in the form of images and tables. The Table 3 shows the performance measures of Decision Tree and Artificial Neural Network for prediction of body weight for boys, girls and both. From the Table 3 it is clear that, Decision Tree regressor has low accuracy in comparison to Artificial Neural Network regressor. Similar results could be interpreted for the measures such as , and .

Table 3 Performance comparison of Decision Tree and Artificial Neural Network for prediction of body weight

	ML	MAE	MSE	RMSE
Boy	DT	0.15	0.076	0.275
Girl		0.392	0.161	0.401
Total		0.268	0.115	0.339
Boy	ANN	0.102	0.036	0.186
Girl		0.116	0.078	0.273
Total		0.185	0.052	0.229

Figure 3(A) shows the regression performance of ANN on test data. The Figure 3(A) (a) represents the regression plot, and the Figure 3(A) (b) represents the residual plot. Since every sample point in the Figure 3(A) (a) is in line with the diagonal reference line, it is obvious that the ANN regressor can accurately estimate body weight. Additionally, it can be seen from the residual plot in Figure 3(A) (b) that the prediction's divergence from the target body weights is modest and falls within the range of -0.5 to 0.5. This demonstrates how important it is that most projections are about similar to actual body weights.

The effectiveness of Decision Tree's regression on test data is shown in Figure 3(B). The regression plot is shown in Figure 3(B)(a), and the residual plot is shown in Figure 3(B) (b). It is evident from Figure 3(B) (a) that the ANN regressor can accurately estimate body weight because all of the sample points are lined up with the diagonal reference line. The divergence of the forecast from the target body weights is modest within the range of -1 to 1, as seen from the Figure 3(B) (b) residual plot. This demonstrates how important it is that while the majority of predictions are roughly comparable to the real body weights, they are less so when compared to the ANN regressor.



(A) (B)

Fig. 3 (A)Regression performance of ANN on test data. The subfigures represent: (a) Regression plot (b) Residual plot.

(B) Regression performance of DT on test data. The subfigures represent: (a) Regression plot (b) Residual plot

Linear model Poly1:
f(x) = p1*x + p2
Coefficients (with 95% confidence bounds):
p1 = 0.9721 (0.9618, 0.9824)
p2 = 0.2786 (0.1721, 0.3852)

Goodness of fit:
SSE: 31.43
R-square: 0.9829
Adjusted R-square: 0.9829
RMSE: 0.2293

(a)

Linear model Poly1:
f(x) = p1*x + p2
Coefficients (with 95% confidence bounds):
p1 = 0.988 (0.9724, 1.004)
p2 = 0.1037 (-0.05849, 0.2658)

Goodness of fit:
SSE: 69.02
R-square: 0.9627
Adjusted R-square: 0.9626
RMSE: 0.3397

(b)

Fig. 4 Regression performance of ANN and DT on test data

The Figure 4 shows the regression analysis of both the ANN and DT regressor. The goodness of fit performances is presented in terms of Sum of Square Error (SSE), R-square, Adjusted R-square and RMSE. The SSE for ANN is 31.43 while for DT it is 69. This represents that, the ANN is predicting the body weight better as compared to the DT regressor. Also, similar interpretation could be observed for R-square, adjusted R-square and RMSE showing the significance of ANN. The obtained results suggest that socioeconomic factor and gender could be considered as one of the factors along with the past body weight information for prediction of body weight.

Conclusion

Effect of maternal education and socio-economic status on growth of neonates is well known in the literature. However, predicting child's body weight using past weight information and socioeconomic factors is limited in research. Also, presently, ML techniques are applied to build such relationship is novel. And from the preliminary findings, to predict the child's body weight in relation with socioeconomic status, the ANN regressor is suggested as compared to the DT regressor.

However, dietary information could be considered as a feature to predict body weight, which could be a future research direction. As the present research is based on data collected from Ganjam district, Odisha, India, several other data can also be gathered and the proposed approach's global significance could be claimed.

Ethical clearance: Written informed consent from legally authorized representatives/parents/guardians are obtained from the participants of the study.

Source of funding: Self

Conflict of Interest: Nil

Reference

1. Swaminathan S, Hemalatha R, Pandey A, Kassebaum NJ, Laxmaiah A, Longvah T, et al. The burden of child and maternal malnutrition and trends in its indicators in the states of India: the Global Burden of Disease Study 1990{2017. The Lancet Child & Adolescent Health. 2019;3(12):855-870.
2. Huang C, Phillips MR, Zhang Y, Zhang J, Shi Q, Song Z, et al. Malnutrition in early life and adult mental health: evidence from a natural experiment. Social science & medicine. 2013; 97:259-266.
3. Alamy M, Bengelloun WA. Malnutrition and brain development: an analysis of the effects of inadequate diet during different stages of life in rat. Neuroscience & Biobehavioral Reviews. 2012;36(6):1463-1480.
4. Guerrant RL, Orifia RB, Moore SR, Orifia MO, Lima AA. Malnutrition as an enteric infectious disease with long-term effects on child development. Nutrition reviews. 2008;66(9):487-505.

5. Narayan J, John D, Ramadas N. Malnutrition in India: status and government initiatives. Journal of public health policy. 2019;40(1):126-141.

6. Panagariya A. Does India really suffer from worse child malnutrition than sub-Saharan Africa? Economic and Political Weekly. 2013;p. 98-111.

7. Vadlamudi S. The Impacts of Machine Learning in Financial Crisis Prediction. Asian Business Review. 2020;10(3):171-176.

8. Patro RN, Subudhi S, Biswal PK, Dell'Acqua F, Sahoo HK. Conditional nearest regularized subspace classifiers: a fast classification approach for HSI. International Journal of Remote Sensing. 2019;40(24):9279-9303.

9. Patro RN, Subudhi S, Biswal PK, Dell'Acqua F. Dictionary-based classifiers for exploiting feature sequence information and their application to hyperspectral remotely sensed data. International Journal of Remote Sensing. 2019;40(13):4996-5024.

10. Dora C, Patro R, Rout S, Biswal P, Biswal B. Adaptive SSA based muscle artifact removal from single channel EEG using neural network regressor. IRBM. 2021;42(5):324-333.

11. Srivastava S, Sharma L, Sharma V, Kumar A, Darbari H. Prediction of diabetes using artificial neural network approach. In: Engineering Vibration, Communication and Information Processing. Springer; 2019. p. 679-687.

12. Etikan I, Caglar MK. Prediction methods for babies' birth weight using linear and non- linear regression analysis. Technology and Health Care. 2005;13(2):131-135.

13. Desiani A, Primartha R, Arhami M, Orsalan O. Naive Bayes classifier for infant weight prediction of hypertension mother. In: Journal of Physics: Conference Series. vol. 1282. IOP Publishing; 2019. p. 012005.

14. Hussain Z, Borah MD. Nutritional status prediction in neonate using machine learning techniques: a comparative study. In: International Conference on Machine Learning, Image Processing, Network Security and Data Sciences. Springer; 2020. p. 69- 83.

15. Al-Shawwa MO, Abu-Naser SS. Predicting Birth Weight Using Artificial Neural Network. 2019;.

16. Senthilkumar D, Paulraj S. Prediction of low birth weight infants and its risk factors using data mining techniques. In: Proceedings of the 2015 international conference on industrial engineering and operations management; 2015. p. 186-194.

17. Celik S, Y_lmaz O. Comparison of different data mining algorithms for prediction of body weight from several morphological measurements in dogs. Journal of Animal and Plant Sciences. 2017;.

18. Wang Z, Shadpour S, Chan E, Rotondo V, Wood KM, Tulpan D. ASAS-NANP SYMPOSIUM: Applications of machine learning for livestock body weight prediction from digital images. Journal of Animal Science. 2021;99(2):skab022.

19. Thangamani D, Sudha P. Identification of malnutrition with use of supervised datamining techniques{decision trees and artificial neural networks. Int J Eng Comput Sci. 2014;3(09).

20. Dora C, Patro R, Rout S, Biswal P, Biswal B. Adaptive SSA Based Muscle Artifact Removal from Single Channel EEG using Neural Network Regressor. IRBM. 2020;.

21. Tan LP, Wong KY. A neural network approach for predicting manufacturing performance using knowledge management metrics. Cybernetics and Systems. 2017;48(4):348-364.

22. Wani RT. Socioeconomic status scales modified Kuppuswamy and Udai Pareekh's scale updated for 2019. Journal of family medicine and primary care. 2019;8(6):1846.