YIDDISHVol 10Issue 1Jan 2020
ISSN NO: 0364-4308

MACHINE LEARNING APPROACH FOR EARLY PREDICTION OF LOW BIRTH WEIGHT (LBW) CASES

Mrs. S.Lavanyaakki , Mr. Likith Sai Bikkinaya

Assistant Professor in the department of CSE at DVR & DR HS MIC COLLEGE OF TECHNOLOGY (Autonomous), Kanchikacherla (NTR Dist, AP)

ABSTRACT_ Low birth weight (LBW) is a medical indicator that something is wrong with a baby. Low birth weight has been associated to a number of negative health outcomes, including infant death. The health of the mother during pregnancy has been linked to a healthier baby, according to a number of studies. In this study, we use machine learning techniques to the health indicators of pregnant women in order to provide early predictions about the likelihood of LBW. By using supervised machine learning for LBW detection, the forecasting issue has been recast as a binary machine classification problem between the LBW and NOT-LBW classes. The suggested model's results exceeded expectations. Decision criteria for predictive health care in smart cities were developed using data from India's health care system. The decision model is being used to create a screening tool for use in Obstetrics and Gynecology (OBG).

1.INTRODUCTION

World Health Organization's Maternal Health and Safe Motherhood Program, 1992. Low Birth Weight. A yearly increase of 12% is predicted. Nearly 39% of power is used for cooling, 45% is used to operate the Data Innovation (IT) framework, and 13% is used for lighting. The high consumption rates result in substantial expenses for the companies. Low birth weight and premature birth are still major public health concerns across the world. Low-birth-weight babies have a far higher risk of developing health problems and dying before their classmates who were born at a normal weight. The percentage of total infant mortality attributable to causes other thanchildren under the age of five dying all around the globe. Babies are considered to have a low birth weight (LBW) if their birth weight is less than 2500 grams. There is widespread agreement that LBW poses a serious threat to public health. LBW include both preterm infants and full-term infants who are extremely tiny owing to intrauterine growth retardation. Infant mortality rates are much higher for low birth weight (LBW) newborns compared to those of infants born at a normal birth weight (NBW). In light of the dire circumstances of the time, this anomaly is causing widespread concern throughout the globe.

long haul issues, for example, improvement problems, neurosensory results, wellbeing results including Type 2 diabetes, cerebral stroke, hypertension and different problems that LBW children are inclined to. Concentrates in 2013 showed that out of the 22 million babies around 16% were low birth weight cases worldwide. This is a major issue in developing nations, particularly in India, which accounts for about 30% of all LBW cases worldwide.

Countless investigations all over the planet demonstrate solid between maternal wellbeing and effect on birth weight of infants. It is a common assumption that dedicated medical care during pregnancy can significantly reduce LBW. Predictions are based on careful examination of the risk factors that can be easily assessed using basic methods in our approach throughout the pregnancy. Early discovery can help in forestalling the possibilities of LBW and furthermore to advance a few suggestions under some mediation components

2.LITERATURE SURVEY

[1] Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ. 1987; 65(5):663-737. PMID: 3322602;PMCID: PMC2491072.

The existence and magnitude of a causal effect on birth weight, gestational age, and prematurity and intrauterine growth retardation were determined by a set of methodological

standards. In developed countries, the most important factor was cigarette smoking, followed by nutrition and pre-pregnancy weight. In developing countries the major determinants were racial origin, nutrition, low pre-pregnancy weight, short maternal stature, and malaria. Pre-pregnancy weight, prior premature birth or miscarriage, diethylstilbestrol exposure and smoking were major determinants of gestational duration, but the majority of prematurity was unexplained in both developed and developing countries.

[2] Vega J, Sáez G, Smith M, Agurto M, Morris NM. Factores de riesgo para bajo peso al nacer y retardo de crecimiento intrauterino en Santiago de Chile [Risk factors for low birth weight and intrauterine growth retardation in Santiago, Chile]. Rev Med Chil. 1993 Oct; 121(10):1210-9. Spanish. PMID: 8191127.

An epidemiologic case-control study to ascertain the determinants of low birth weight was carried out in Santiago, Chile, from January to December 1989. The cases were defined as live births < 2500 g. The controls were live births > or = 2500 g of birth weight. All cases and a random sample (1:1) of controls were selected among 8,254 singleton births occurring at the El Salvador Hospital in the Eastern area of Santiago. These deliveries represented 50% of institutional deliveries in the area. Home deliveries (2%) and private hospital deliveries were not included in the study. Information was obtained from hospital medical records by six trained medical students. Some information could not be obtained from the hospital medical records. Thus the second step in data collection was the tracking of all the selected subjects to their referring neighborhood health centers.

[3] Mavalankar DV, Trivedi CC, Gray RH. Maternal weight, height and risk of poor pregnancy outcome in Ahmedabad, India. Indian Pediatr. 1994 Oct; 31(10):1205-12. PMID: 7875780.

This paper explores the relationships between maternal weight, height and poor pregnancy outcome using a data set from a case-control study of low birth weight (LBW) and perinatal mortality in Ahmedabad, India. Maternal height and weights were compared between mothers of 611 perinatal deaths, 644 preterm- LBW, and 1465 normal birth weight controls as well as 617 small-for-gestational age (SGA) and 1851 appropriate-for-gestational-age (AGA) births. Weight and height were much lower in this population compared to western standards. Low weight and height were associated with increased risk of perinatal death, prematurity and SGA. After adjusting for confounders, maternal weight remained significantly associated with poor pregnancy outcomes, whereas height was only weakly associated. Attributable risk estimates show that low weight is a much more important contributor to poor outcome than low height. Improvement in maternal nutritional status

could lead to substantial improvement in birth outcome in this population

3.PROPOSED SYSTEM

We utilise supervised machine learning methods such as XGBoost Classifier, Random Forest Classifier, Support Vector Classifier, and Decision Tree Classifier in the proposed system to detect low birth weight newborns..

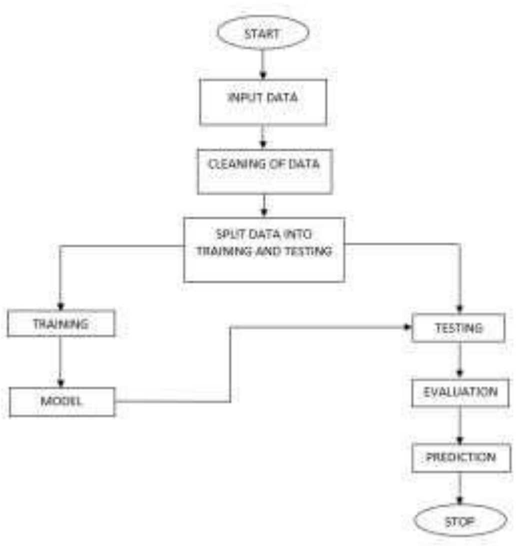

FIGURE 1Flow chart

IMPLEMENTATIONXGBoost:

XGBoost is an algorithm that has recently been dominating applied machine learning and Kaggle competitions for structured or tabular data. XGBoost is an implementation of gradient boosted decision trees designed for speed and performance.

XGBoost is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework. In prediction problems involving unstructured data (images, text, etc.) artificial neural networks tend to

outperform all other algorithms or frameworks. However, when it comes to small-to-medium structured/tabular data, decision tree basedalgorithms are considered best-in-class right now.

Bagging: Now imagine instead of a single interviewer, now there is an interview panel where each interviewer has a vote. Bagging or bootstrap aggregating involves combining inputs from all interviewers for the finaldecision through a democratic voting process.

XGBoost and Gradient Boosting Machines (GBMs) are both ensemble tree methods that apply the principle of boosting weak learners (CARTs generally) using the gradient descent architecture. However, XGBoost improves upon the base GBM framework through systems optimization and algorithmic enhancements.

Random Forest:

First, Random Forest algorithm is a supervised classification algorithm. We can see it from its name, which is to create a forest by some way and make it random. There is a direct relationship between the number of trees in the forest and the results it can get: the larger the number of trees, the more accurate the result. But one thing to note is that creating the forest is not the same as constructing the decisionwith information gain or gain index approach.

The author gives four advantages to illustrate why we use Random Forest algorithm. The one

mentioned repeatedly by the author is that it can be used for both classification and regression tasks. Overfitting is one critical problem that may make the results worse, but for Random Forest algorithm, if there are enough trees in the forest, the classifier won’t overfit the model. The third advantage is the classifier of Random Forest can handle missing values, and the last advantage is that the Random Forest classifier can be modeled for categorical values.

There are two stages in Random Forest algorithm, one is random forest creation, the other is to make a prediction from the random forest classifier created in the first stage.

STEPS:

1. Randomly select “K” features from total “m” features where k << m
2. Among the “K” features, calculate the node“d” using the best split point
3. Split the node into daughter nodes using the best split
4. Repeat the a to c steps until “l” number of nodes has been reached
5. Build forest by repeating steps a to d for “n” number times to create “n” number of trees

Decision Trees:

A tree has many analogies in real life, and turns out that it has influenced a wide area of machine learning, covering both classification and regression. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. As the name goes, it uses a tree-like model of decisions. Though a commonly used tool in data mining for deriving a strategy to reach a particular goal.

A decision tree is drawn upside down with its root at the top. In the image on the left, the bold text in black represents a condition/internal node, based on which the tree splits into branches/ edges. The end of the branch that doesn’t split anymore is the decision/leaf, in this case, whether the passenger died or survived, represented as red and green text respectively.

Although, a real dataset will have a lot more features and this will just be a branch in a much bigger tree, but you can’t ignore the simplicity of this algorithm. The feature importance is clear and relations can be viewed easily. This methodology is more commonly known as learning decision tree from data and above tree is called Classification tree as the target is to classify passenger as survived or died. Regression trees are represented in the same manner, just they predict continuous values likeprice of a house. In general, Decision Tree algorithms are referred to as CART or Classification and Regression Trees.

So, what is actually going on in the background? Growing a tree involves deciding on which features to choose and what conditions to use for splitting, along with knowing when to stop. As a tree generally grows arbitrarily, you will need to trim it down for it to look beautiful. Let’s start with a common technique used for splitting.

Support Vector Machine:

A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model sets of labeled training data for eachcategory, they're able to categorize new text.

So you're working on a text classification problem. You're refining your training data, and maybe you've even tried stuff out using Naive Bayes. But now you're feeling confident in your dataset, and want to take it one step further. Enter Support Vector Machines (SVM): a fast and dependable classification algorithm that performs very well with a limited amount of data to analyze.

Perhaps you have dug a bit deeper, and ran into terms like linearly separable, kernel trick and kernel functions. But fear not! The idea behind the SVM algorithm is simple, and applying it to natural language classification doesn't require most of the complicated stuff.

Steps for implementation:

- Import the dataset.
- Explore the data to figure out what theylook like.
- Pre-process the data.
- Split the data into attributes and labels.
- Divide the data into training and testingsets.
- Train the SVM algorithm.
- Make some predictions

4.RESULTS AND DISCUSSION

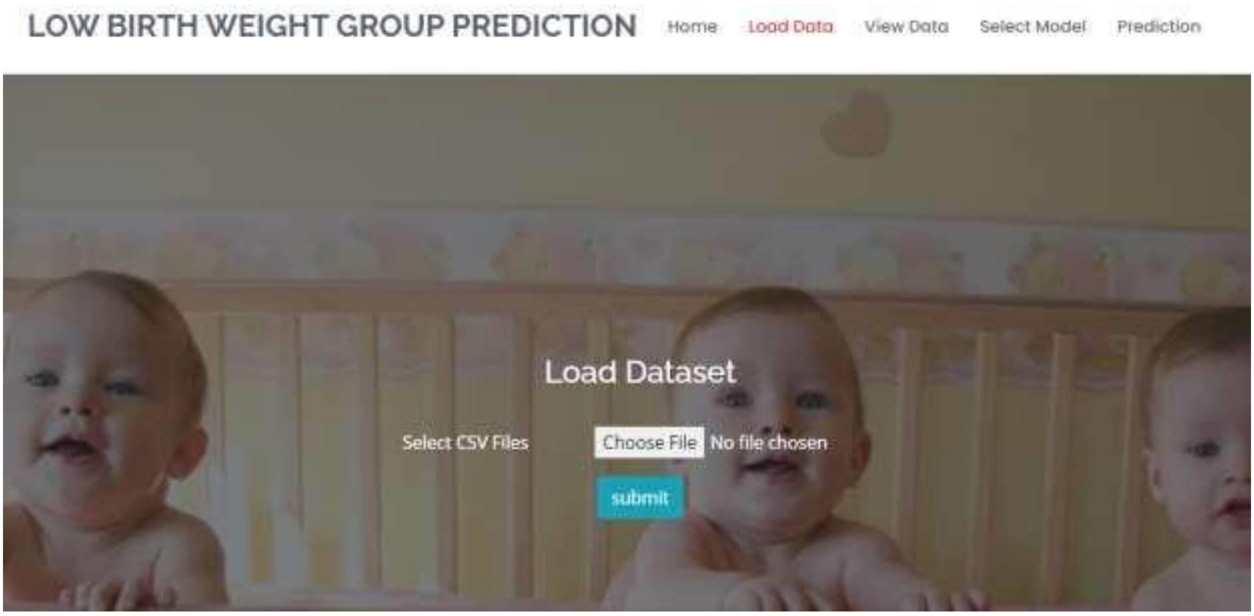

Fig 2:Upload Dataset Page

S/N	community	age	weight1	history	HB	IFA	BP1	education	res	result
1	1.0	26.0	37.0	1.0	5.9	1.0	1.4444444440000002	5.0	1.0	0.0
2	1.0	21.0	42.0	1.0	9.2	1.0	1.375	5.0	1.0	0.0
3	1.0	21.0	47.136364	1.0	8.8	1.0	1.5	5.0	1.0	0.0
4	1.0	21.0	47.136364	1.0	9.2	1.0	2.125	5.0	1.0	0.0
5	1.0	21.0	47.136364	1.0	8.0	1.0	1.375	5.0	1.0	0.0
6	1.0	24.0	33.0	1.0	9.3	1.0	1.571	5.0	1.0	0.0
7	1.0	26.0	35.0	1.0	9.2	1.0	1.571428571	5.0	1.0	0.0
8	4.0	26.0	31.0	1.0	9.076922999999999	1.0	1.625	5.0	1.0	0.0
9	3.0	21.0	47.136364	1.0	11.0	1.0	1.375	5.0	1.0	0.0
10	1.0	22.0	30.0	1.0	9.0	1.0	1.482	5.0	1.0	0.0
11	4.0	17.0	30.0	1.0	9.0	0.0	1.375	5.0	1.0	0.0
12	3.0	35.0	54.0	1.0	9.9	1.0	1.571428571	5.0	1.0	0.0

Fig 3:DATA VIEWING PAGE:

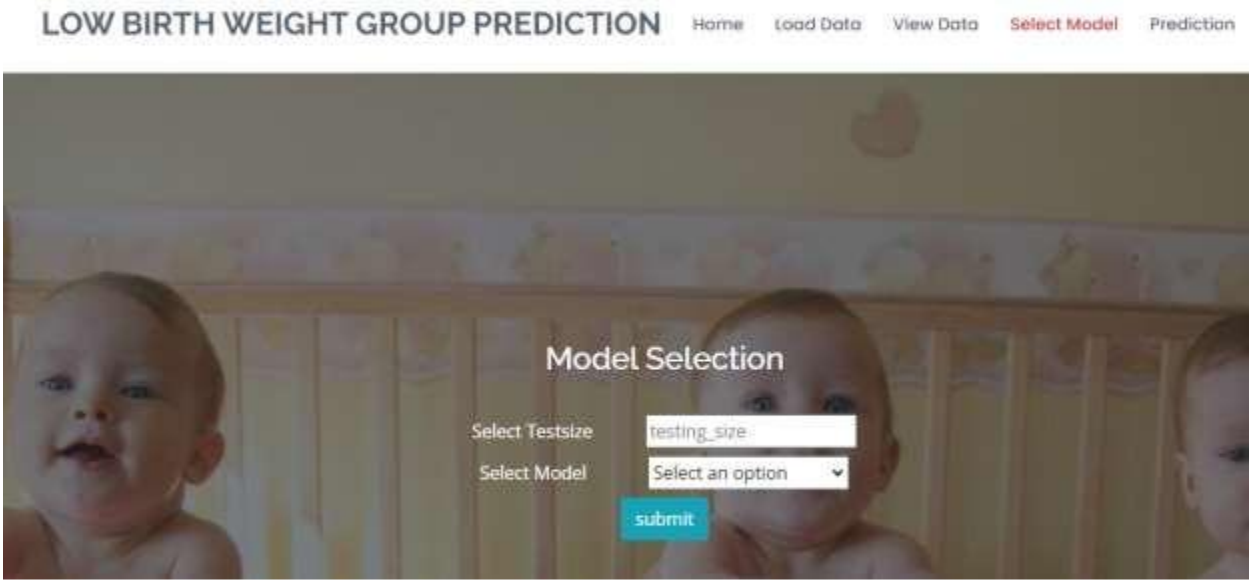

Fig 4:MODEL SELECTION PAGE:

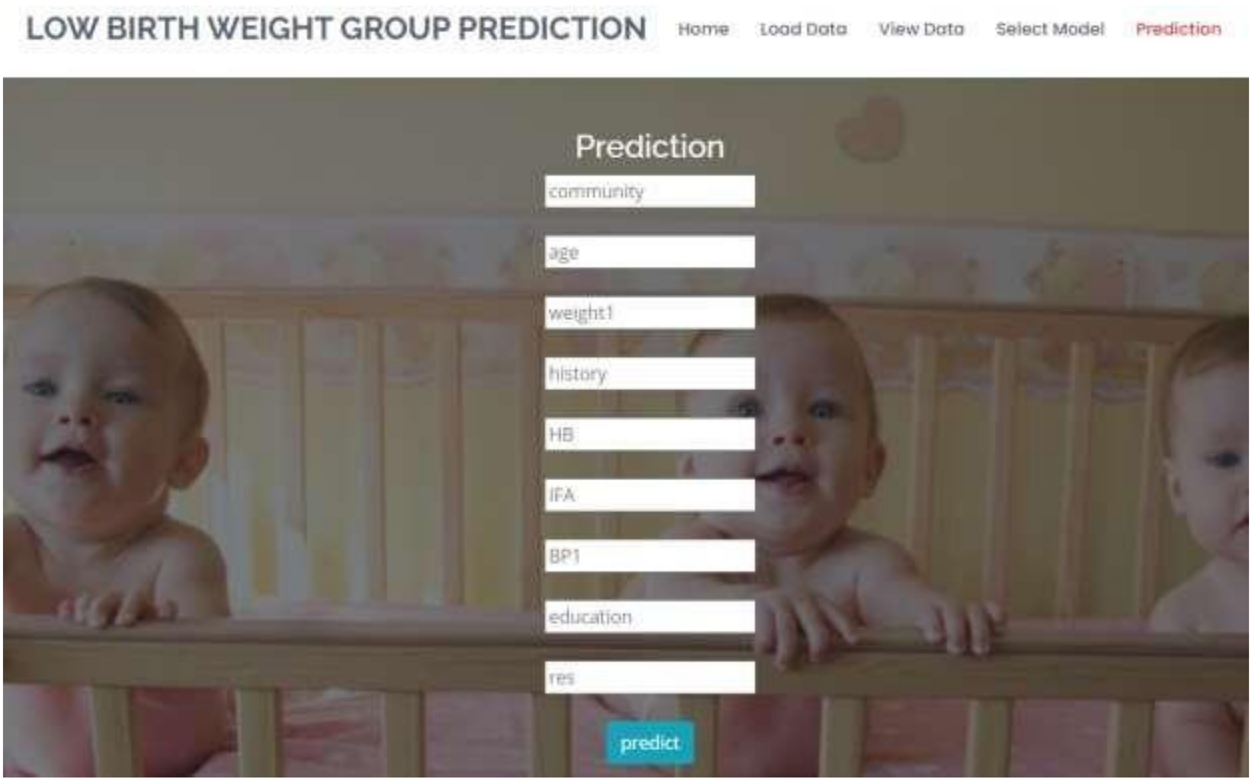

Fig 5:PREDICTION PAGE:

5.CONCLUSION

In this application, we successfully constructed ML models to assess whether or not the baby belongs to the Low Birth Weight category. This is created in an easy-to-use environment with Flask and Python programming. We discovered that the Decision Tree Classifier outperforms the XGBoost Classifier, Random Forest Classifier, Decision Tree Classifier, and Support Vector Classifier in terms ofaccuracy..

REFERENCES

[1] World Health Organization-1992, International statistical classification of diseases and related health problems,Tenthrevision, Geneva, World HealthOrganization.

[2] Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. Bull World Health Organ. 1987; 65(5):663-737. PMID: 3322602;PMCID: PMC2491072.

[3] Vega J, Sáez G, Smith M, Agurto M, Morris NM. Factores de riesgo para bajo peso al nacer y retardo de crecimiento intrauterino en Santiago de Chile [Risk factors for low birth weight and intrauterine growth retardation in Santiago, Chile]. Rev Med Chil. 1993 Oct; 121(10):1210-9. Spanish. PMID: 8191127.

[4] Mavalankar DV, Trivedi CC, Gray RH. Maternal weight, height and risk of poor pregnancy outcome in Ahmedabad, India.
Indian Pediatr. 1994 Oct; 31(10):1205-12.
PMID: 7875780.

[5] Bosetti C, Nieuwenhuijsen MJ, Gallus S, Cipriani S, La Vecchia C, Parazzini F. Ambient particulate matter and preterm birth or birth weight: a review of the literature. Arch Toxicol. 2010 Jun; 84(6):447-60. Doi: 10.1007/s00204-010- 0514-z. Epub 2010 Feb 6. PMID: 20140425.

[6] United nations Children's Fund and World Health Organization2004, Low Birth Weight: Country, regional and globalestimates, New York, UNICEF.

[7] J.S. Deshmukh, D.D. Motghare, S.P. Zodpey and S.K. Wadhva1998, Low Birth Weight

And Associated Maternal Factors in an Urban Area, Indian Pediatrics,Volume 35, Page 33-36

[8] J.S. Deshmukh, D.D. Motghare, S.P. Zodpey and S.K. Wadhva1998, Low Birth Weight And Associated Maternal Factors in an Urban Area, Indian Pediatrics,Volume 35, Page 33-36

[9] Aparajita Dasgupta, Rivu Basu, “Determinants of low birth weight in a Blowk of Hooghly, West Bengal: A multivariate analysis,” International Journal of Biological & Medical Research,2(4), 2011, pp.838-842.

[10] Bellazzi R, Zupan B., “Towards knowledge-based gene expression data mining”, J Biomed Inform. 2007; 40(6),pp.787-802

completed BSC in TRIVENI DEGREE COLLEGE From ADIKAVI NANNAYA UNIVERSITY. His areas of interests are Networks, Machine Learning and Cloud Computing.