

CO544 TUTORIAL 1

RANAGE R.D.P.R.

E/19/310

SEMESTER 6

PAPER ID 4

09/04/2024

Paper ID: 4

Performance Analysis of Supervised Learning Classifiers for the Prediction of Child Birth Weight

1. What is the problem being solved? Be concise.

Using supervised learning classifiers, the problem of predicting a child's birth weight is being solved. In order to create a predictive model that can categorize newborns into groups like Normal Birth Weight (NB), Low Birth Weight (LB), and High Birth Weight (HB) based on specific parameters extracted from the data set, the existing data on low birth weight (LBW) in the Ampara region of Sri Lanka must be used.

2. Is it a supervised or unsupervised learning problem (Answer in Yes/No)? If it is a supervised learning problem, what are labels? If it is an unsupervised learning problem what is the output?

Yes, it is a supervised learning problem.

Labels: Normal Birth Weight (NB)
Low Birth Weight (LB)
High Birth Weight (HB)

3. What is/are the reason/s to consider the particular Machine Learning method to solve this problem?

Because of its great accuracy and short processing time, the C4.5 decision tree classifier was selected as the specific machine learning technique for predicting infant birth weight. Predicting birth weight categories is one of the categorization problems that the C4.5 algorithm is well-suited for because of its reputation for efficiently building decision trees that can manage both numerical and categorical data. Furthermore, C4.5 is preferred because of its capacity to manage missing values in the data set during pre-processing, an essential step in guaranteeing the prediction model's accuracy. The selection of C4.5 for this predictive modeling assignment was based on its ability to reliably classify babies into the relevant weight groups while keeping a reasonably low processing time.

4. What metrics have been used to evaluate the proposed solution?

Accuracy: This metric measures how well the predictive model classifies newborns into the correct weight categories (Normal Birth Weight, Low Birth Weight, High Birth Weight). A higher accuracy percentage indicates that the model is making more correct predictions, which is essential for ensuring the reliability and effectiveness of the predictive model.

Time Complexity: Time complexity refers to the computational efficiency of the decision tree classifiers, specifically how long it takes to construct the decision tree model. Lower time complexity values indicate that the model can be built quickly, which is important for practical applications where efficiency and speed are crucial factors.